BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

OPIC
Office de la propriété intellectuelle du Canada

CIPO
Canadian Intellectual Property Office

Ottawa Hull K1A 0C9

(21) (A1) 2,157,596
(86) 1994/03/09
(43) 1994/09/29

(51) Int.Cl.[6] C12N 15/29; C12N 5/10; C12N 5/08; C07K 16/16; C
14/415; C07K 14/725; A61K 39/35; G01N 33/564

(19)(CA) **APPLICATION FOR CANADIAN PATENT** (12)

(54) T Cell Epitopes of Ryegrass Pollen Allergen

(72) Griffith, Irwin J. - U.S.A. ;
Kuo, Mei-Chang - U.S.A. ;
Luqman, Mohammad - U.S.A. ;
Powers, Stephen Palmer - U.S.A. ;

(71) Immulogic Pharmaceutical Corporation - U.S.A. ;

(30) (US) 08/031,001 1993/03/12

(57) 71 Claims

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C07K 7/08, 15/10, 15/28
C12N 15/29, 15/70, 1/21, A61K 37/02

A3

(11) International Publication Number: WO 94/21675

(43) International Publication Date: 29 September 1994 (29.09.94)

(21) International Application Number: PCT/US94/02537

(22) International Filing Date: 9 March 1994 (09.03.94)

(30) Priority Data:
08/031,001 12 March 1993 (12.03.93) US

(71) Applicant: IMMULOGIC PHARMACEUTICAL CORPORATION [US/US]; 610 lincoln Street, Waltham, MA 02154 (US).

(72) Inventors: GRIFFITH, Irwin, J.; 13 Southwick Road, North Reading, MA 01864 (US). KUO, Mei-chang; 5 Cox Road, Winchester, MA 01890 (US). LUQMAN, Mohammad; 13 Carriage Drive, Acton, MA 01720 (US). POWERS, Stephen, Palmer; 2008 Stearns Hill Road, Waltham, MA 02154 (US).

(74) Agents: CHANNING, Stacey, L. et al.; Immulogic Pharmaceutical Corporation, 610 Lincoln Street, Waltham, MA 02154 (US).

(81) Designated States: AU, BB, BG, BR, BY, CA, CN, CZ, FI, HU, JP, KP, KR, KZ, LK, LV, MG, MN, MW, NO, NZ, PL, RO, RU, SD, SK, UA, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
10 November 1994 (10.11.94)

2157596

(54) Title: T CELL EPITOPES OF RYEGRASS POLLEN ALLERGEN

(57) Abstract

The present invention provides isolated peptides of *Lol p* I, a major protein allergen of the species *Lolium perenne*. Peptides within the scope of the invention comprise at least one T cell epitope, or preferably at least two T cell epitopes of a protein allergen of *Lol p* I. The invention also provides modified peptides having similar or enhanced therapeutic or diagnostic properties as the corresponding, naturally-occurring allergen or portion thereof, but having additional properties, e.g., reduced side effects. The invention further provides nucleic acid sequences coding for peptides of the invention. Methods of treatment and diagnosis of sensitivity to *Lol p* I or an allergen immunologically related to *Lol p* I in an individual (such as *Dac g* I, *Poa p* I, or *Phl p* I) also are provided. Compositions for therapeutic, diagnostic or reagent uses comprising one or more peptides of the invention are also provided.

2157596

Claims

1. An isolated peptide of *Lol p* I or an isolated portion thereof, said peptide or portion thereof comprising at least one T cell epitope of *Lol p* I, said peptide comprising an amino acid sequence selected from the group consisting of: LPI-1 (SEQ ID NO: 4).1, LPI-2 (SEQ ID NO: 5), LPI-3 (SEQ ID NO: 6), LPI-4 (SEQ ID NO: 7), LPI-4.1 (SEQ ID NO: 8), LPI-8 (SEQ ID NO: 12), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-13 (SEQ ID NO: 19), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-19 (SEQ ID NO: 26), LPI-22 (SEQ ID NO: 29) and LPI-23 (SEQ ID NO: 30), all as shown in Fig. 2.

2. An isolated peptide of *Lol p* I or an isolated portion thereof, said peptide or portion thereof comprising at least one T cell epitope of *Lol p* I, said peptide having an amino acid sequence selected from the group consisting of: LPI-16.2 (SEQ ID NO: 31), LPI-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID NO: 33), LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID NO: 35), LPI-16.7 (SEQ ID NO: 36), LPI-16.9 (SEQ ID NO: 37), LPI-16.10 (SEQ ID NO: 38), LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-18.8 (SEQ ID NO: 42), LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), LPI-20.6 (SEQ ID NO: 47), LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49), and LPI-23.4 (SEQ ID NO: 50), all as shown in Fig. 4.

3. An isolated peptide or portion thereof according to claim 1, wherein said portion of a peptide has a mean T cell stimulation index approximately equivalent to or greater than the mean T cell stimulation index of the corresponding peptide shown in Fig. 3.

4. An isolated peptide or portion thereof of claim 1 or 2 which comprises at least two T cell epitopes.

5. An isolated peptide or portion thereof of claim 1 or 2 which induces T cell nonresponsiveness or modifies the lymphokine secretion profile of appropriate T cell subpopulations.

6. An isolated peptide or portion thereof of claim 1 or 2 which, when administered to an individual sensitive to an allergen of the family, Poacea induces T cell anergy or modifies the lymphokine secretion profile of approprate T cell populations.

7. A portion of an isolated peptide of claim 1 or 2 which has a mean T cell stimulation index of at least 3.5.

8. An isolated peptide or a portion thereof of claim 1 or 2 which does not bind immunoglobulin E specific for *Lol p* I in a substantial percentage of individuals sensitive to *Lol p I,* or if binding of the peptide or portion thereof to said immunoglobulin E occurs, such binding does not result in release of mediators from mast cells or basophils in a substantial percentage of individuals sensitive to *Lol p* I.

9. An isolated peptide of claim 1 or 2 which binds immunoglobulin E to a substantially lesser extent than purified native *Lol p* I binds immunoglobulin E.

10. An isolated peptide or portion thereof of claim 1 or 2 which, when administered to an individual sensitive to *Lol p* I allergen, modifies the allergic response of the individual to ryegrass pollen allergen.

11. An isolated peptide or portion thereof of claim 1 or 2 which, when administered to an individual sensitive to an allergen of the family Poacea, modifies the allergic response of the individual to said allergen.

12. A portion of an isolated peptide of claim 1 or 2 wherein said portion comprises at least 15 amino acid residues.

13. An isolated nucleic acid having a sequence encoding all or a portion of a peptide of claim 1 or 2.

14. A functional equivalent of a nucleic acid sequence encoding all or a portion of a peptide of claim 1 or 2.

15. An isolated peptide that is immunologically cross-reactive with T cells reactive with a peptide of claim 1 or 2.

16. An isolated peptide or portion thereof of *Lol p* I, said peptide or portion thereof comprising at least one T cell epitope of *Lol p* I, said peptide having a positivity index of at least about 100 and mean T cell stimulation index of at least about 3.0 determined in a population of individuals sensitive to said protein allergen.

17. An isolated peptide or portion thereof of claim 16 wherein said population of individuals is at least thirty individuals.

18. An isolated peptide or portion thereof of claim 17 wherein said population of individuals is at least thirty-five individuals.

19. An isolated peptide or portion thereof of claim 17 wherein said mean T cell stimulation index is at least about 4.0.

20. An isolated peptide or portion thereof of claim 17 wherein said mean T cell stimulation index is at least about 6.0.

21. A peptide or portion thereof of claim 17 wherein said peptide is selected from the group consisting of: LPI-2 (SEQ ID NO: 5), LPI-11 (SEQ ID NO: 15), LPI-13 (SEQ ID NO: 19), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-22 (SEQ ID NO: 29) and LPI-23 (SEQ ID NO: 30).

22. An isolated peptide of *Lol p* I, or a portion thereof wherein said peptide is selected from the group consisting of: LPI-1.1 (SEQ ID NO: 4), LPI-2 (SEQ ID NO: 5), LPI-3 (SEQ ID NO: 6), LPI-4 (SEQ ID NO: 7), LPI-4.1 (SEQ ID NO: 8), LPI-8 (SEQ ID NO: 12), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-13 (SEQ ID NO: 19), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-19 (SEQ ID NO: 26), LPI-22 (SEQ ID NO: 29), LPI-23 (SEQ ID NO: 30), LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-

18.8 (SEQ ID NO: 42), LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), LPI-20.6 (SEQ ID NO: 47), LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49), and LPI-23.4 (SEQ ID NO: 50) or portion thereof.

23. A modified peptide or a modified portion of a peptide of claim 22.

24. A modified peptide of claim 23 wherein said peptide is selected from the group consisting of: LPI-16.2 (SEQ ID NO: 31), LPI-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID NO: 33), LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID NO: 35), LPI-16.7 (SEQ ID NO: 36), LPI-16.9 (SEQ ID NO: 37), and LPI-16.10 (SEQ ID NO: 38), all as shown in Fig. 4.

25. A modified peptide or a modified portion of a peptide of claim 23 or 24 which does not bind immunoglobulin E specific for *Lol p* I in a substantial percentage of individuals sensitive to *Lol p* I, or if binding of the peptide or portion thereof to said immunoglobulin E occurs, such binding does not result in release of mediators from mast cells or basophils in a substantial percentage of individuals sensitive to *Lol p* I.

26. A modified peptide or a modified portion of a peptide of claim 23 or 24 which modifies, in an individual sensitive to *Lol p* I or an immunologically related allergen, the allergic response of the individual to *Lol p* I allergen or said related allergen.

27. An isolated peptide comprising at least two regions, each region comprising at least one T cell epitope of *Lol p* I, said regions each comprising all or a portion of an amino acid sequence selected from the group consisting of: LPI-1.1 (SEQ ID NO: 4), LPI-2 (SEQ ID NO: 5), LPI-3 (SEQ ID NO: 6), LPI-4 (SEQ ID NO: 7), LPI-4.1 (SEQ ID NO: 8), LPI-8 (SEQ ID NO: 12), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-13 (SEQ ID NO: 19), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), LPI-23 (SEQ ID NO: 30), LPI-16.2 (SEQ ID NO: 31), LPI-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID NO: 33), LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID

NO: 35), LPI-16.7 (SEQ ID NO: 36), LPI-16.9 (SEQ ID NO: 37), LPI-16.10 (SEQ ID NO: 38), LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-18.8 (SEQ ID NO: 42), LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), LPI-20.6 (SEQ ID NO: 47), LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49), and LPI-23.4 (SEQ ID NO: 50).

28. An isolated peptide of claim 27 wherein said regions comprise an amino acid sequence selected from the group consisting of: LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), LPI-23 (SEQ ID NO: 30), LPI-16.2 (SEQ ID NO: 31), LPI-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID NO: 33), LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID NO: 35), LPI-16.7 (SEQ ID NO: 36), LPI-16.9 (SEQ ID NO: 37), LPI-16.10 (SEQ ID NO: 38), LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-18.8 (SEQ ID NO: 42), LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), LPI-20.6 (SEQ ID NO: 47), LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49), and LPI-23.4 (SEQ ID NO: 50), or a portion thereof containing at least two *Lol p* I epitopes.

29. An isolated peptide of *Lol p* I, wherein said peptide comprises a combination of regions selected from the group consisting of:

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), and LPI-11 (SEQ ID NO: 15);

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), and LPI-16 (SEQ ID NO: 22);

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), and LPI-16.1 (SEQ ID NO: 23);

LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), and LPI-16.1 (SEQ ID NO: 23);

LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), and LPI-20 (SEQ ID NO: 27);

LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29) and LPI-23 (SEQ ID NO: 30);

LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), and LPI-20 (SEQ ID NO: 27);

LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);

LPI-18 (SEQ ID NO: 25) and LPI-20 (SEQ ID NO: 27);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27) and LPI-23 (SEQ ID NO: 30);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27) and LPI-16.1 (SEQ ID NO: 23);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30) and LPI-16.1 (SEQ ID NO: 23);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23) and LPI-11 (SEQ ID NO: 15);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23) and LPI-4.1 (SEQ ID NO: 8);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16 (SEQ ID NO: 23).1, LPI-4.1 (SEQ ID NO: 8) and LPI-22 (SEQ ID NO: 29);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23), LPI-11 (SEQ ID NO: 15) and LPI-4.1 (SEQ ID NO: 8);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23), LPI-11 (SEQ ID NO: 15), LPI-4.1 (SEQ ID NO: 8) and LPI-22 (SEQ ID NO: 29);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-16.1 (SEQ ID NO: 23), LPI-22 (SEQ ID NO: 29) and LPI-23 (SEQ ID NO: 30); and
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-16.1 (SEQ ID NO: 23) and LPI-22 (SEQ ID NO: 29).

30. An isolated peptide of *Lol p* I, wherein said peptide comprises a combination of regions selected from the group consisting of:
LPI-16.2 (SEQ ID NO: 31), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.3 (SEQ ID NO: 32), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.4 (SEQ ID NO: 33), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.5 (SEQ ID NO: 34), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.6 (SEQ ID NO: 35), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.7 (SEQ ID NO: 36), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.9 (SEQ ID NO: 37), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30); and
LPI-16.10 (SEQ ID NO: 38), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30).

31. A monoclonal antibody, polyclonal antibody, or immunoreactive fragment thereof specifically reactive with a peptide of claim 1 or 2.

32. An isolated peptide produced in a host cell transformed with the nucleic acid of claim 13.

33. An isolated peptide produced in a host cell transformed with the nucleic acid of claim 14.

34. An isolated nucleic acid having a sequence encoding a peptide of claim 27 or 29.

35. The functional equivalent of an isolated nucleic acid sequence encoding a peptide of claim 27 or 29.

36. An isolated peptide produced in a host cell transformed with the nucleic acid of claim 34.

37. An expression vector comprising a nucleic acid sequence coding for a peptide of claim 1 or 2.

38. An expression vector comprising the functional equivalent of a sequence coding for a peptide of claim 1 or 2.

39. An expression vector comprising a nucleic acid sequence coding for a peptide of claim 27 or 29.

40. An expression vector comprising the functional equivalent of a nucleic acid sequence coding for a peptide of claim 27 or 29.

41. All or a portion of an isolated peptide of *Lol p I*, said peptide or portion thereof comprising at least one T cell epitope of said protein allergen, said peptide having the formula X_n-Y-Z_m, wherein Y is an amino acid sequence selected from the group consisting of: LPI-1 (SEQ ID NO: 3), LPI-1.1 (SEQ ID NO: 4), LPI-2 (SEQ ID NO: 5), LPI-3 (SEQ ID NO: 6), LPI-4 (SEQ ID NO: 7), LPI-4.1 (SEQ ID NO: 8), LPI-5 (SEQ ID NO: 9), LPI-6 (SEQ ID NO: 10), LPI-7 (SEQ ID NO: 11), LPI-8 (SEQ ID NO: 12), LPI-9 (SEQ ID NO: 13), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-12 (SEQ ID NO: 17), LPI-13 (SEQ ID NO: 19), LPI-14 (SEQ ID NO: 20), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-17 (SEQ ID NO: 24), LPI-18 (SEQ ID NO: 25), LPI-19 (SEQ ID NO: 26), LPI-21 (SEQ ID NO: 28), LPI-22 (SEQ

LPI-23.2 (SEQ ID NO: 49), and LPI-23.4 (SEQ ID NO: 50) wherein X_n are amino acid residues contiguous to the amino terminus of Y in the amino acid sequence of said protein allergen, wherein Z_m are amino acid residues contiguous to the carboxy terminus of Y in the amino acid sequence of said protein allergen, wherein n is 0-30 and wherein m is 0-30.

42. A portion of an isolated peptide of claim 40 wherein the portion comprises at least fifteen amino acid residues.

43. A composition comprising at least one isolated peptide or a portion thereof of claim 1 or 2 and a pharmaceutically acceptable carrier or diluent.

44. A composition comprising at least one isolated peptide or portion thereof of claim 23 or 24 and a pharmaceutically acceptable carrier or diluent.

45. A therapeutic composition comprising an isolated peptide or portion thereof of claim 27 or 29 and a pharmaceutically acceptable carrier or diluent.

46. Use of a composition of claim 43 in the manufacture of a medicament for treating sensitivity to *Lol p* I protein allergen or an allergen which is immunologically cross reactive with *Lol p* I protein allergen.

47. Use of a composition of claim 44 in the manufacture of a medicament for treating sensitivity to *Lol p* I protein allergen or an allergen which is immunologically cross reactive with *Lol p* I protein allergen.

48. Use of at least two compositions of claim 43 in the manufacture of a medicament for treating sensitivity to *Lol p* I protein allergen or an allergen which is immunologically cross reactive with *Lol p* I protein allergen.

49. The use of the composition of claim 46 in the manufacture of a medicament for treating sensitivity to *Lol p* I protein allergen, wherein said immunologically cross reactive allergen is *Dac g* I, *Poa p* I or *Phl p* I.

50. A method of detecting sensitivity to *Lol p* I protein allergen or an immunlogically cross reactive allergen in an individual, comprising combining a blood sample obtained from the individual with at least one peptide of claim 1 or 2, *in vitro*, under conditions appropriate for binding of

blood components with the peptide, and determining the extent to which such binding occurs as indicative of sensitivity in the individual to ryegrass pollen allergen or said immunlogically cross-reactive allergen.

51. A method of claim 50 wherein the extent to which binding occurs is determined by assessing T cell function, T cell proliferation or a combination thereof.

52. A composition comprising a pharmaceutically acceptable carrier or diluent and at least two peptides, selected from the group consisting of: LPI-1.1 (SEQ ID NO: 4), LPI-2 (SEQ ID NO: 5), LPI-3 (SEQ ID NO: 6), LPI-4 (SEQ ID NO: 7), LPI-4.1 (SEQ ID NO: 8), LPI-8 (SEQ ID NO: 12), LPI-11 (SEQ ID NO: 15), LPI-13 (SEQ ID NO: 19), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), LPI-23 (SEQ ID NO: 30), LPI-16.2 (SEQ ID NO: 31), LPI-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID NO: 33), LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID NO: 35), LPI-16.7 (SEQ ID NO: 36), LPI-16.9 (SEQ ID NO: 37), LPI-16.10 (SEQ ID NO: 38), LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-18.8 (SEQ ID NO: 42), LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), LPI-20.6 (SEQ ID NO: 47), LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49), and LPI-23.4 (SEQ ID NO: 50) and wherein said composition comprises a sufficient percentage of the T cell epitopes of said protein allergen such that T cells of an individual sensitive to *Lol p* I protein pollen or an immunologically cross reactive allergen, are tolerized to said at least one protein allergen.

53. A composition of claim 45 comprising a combination of peptides selected from the group consisting of:

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), and LPI-11 (SEQ ID NO: 15);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23) and LPI-11 (SEQ ID NO: 15);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23) and LPI-4.1 (SEQ ID NO: 8);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23), LPI-4.1 (SEQ ID NO: 8) and LPI-22 (SEQ ID NO: 29);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23), LPI-11 (SEQ ID NO: 15) and LPI-4.1 (SEQ ID NO: 8);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23), LPI-11 (SEQ ID NO: 15), LPI-4.1 (SEQ ID NO: 8) and LPI-22 (SEQ ID NO: 29);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-16.1 (SEQ ID NO: 23), LPI-22 (SEQ ID NO: 29) and LPI-23 (SEQ ID NO: 30); and
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-16.1 (SEQ ID NO: 23) and LPI-22 (SEQ ID NO: 29).

54. A composition of claim 43 comprising a combination of peptides selected from the group consisting of:
LPI-16.2 (SEQ ID NO: 31), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.3 (SEQ ID NO: 32), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.4 (SEQ ID NO: 33), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.5 (SEQ ID NO: 34), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.6 (SEQ ID NO: 35), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.7 (SEQ ID NO: 36), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);

LPI-16.9 (SEQ ID NO: 37), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30); and
LPI-16.10 (SEQ ID NO: 38), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30).

55. Use of composition of claim 52, 53 or 54 in the manufacture of a medicament for use in treating sensitivity to *Lol p* I allergen or an immunologically cross-reactive allergen.

56. A method of designing antigenic fragments of *Lol p* I, which when administered to ryegrass pollen sensitive individuals in sufficient quantity will modify the individual's allergic reaction to ryegrass pollen comprising the steps of:
(a) recombinantly or synthetically producing peptides of *Lol p* I;
(b) examining said peptides for their ability to influence B cell and/or T cell responses in ryegrass pollen sensitive individuals;
(c) selecting appropriate peptides which contain epitopes recognized by the cells, and
(d) combining epitope-containing regions to include multiple epitopes in one peptide.

57. A method of designing antigenic fragments of *Lol p* I, which when administered to ryegrass pollen sensitive individuals in sufficient quantity will modify the individual's allergic reaction to ryegrass pollen comprising the steps of:
(a) recombinantly or synthetically producing peptides of *Lol p I*;
(b) examining said peptides for their ability to influence B cell and/or T cell responses in ryegrass pollen sensitive individuals; and
(c) selecting appropriate peptides which contain epitopes recognized by the cells.

58. A T cell capable of recognizing a peptide of claim 1 or 2.

59. A receptor of a T cell capable of recognizing a peptide of claim 1 or 2.

60. An isolated nucleic acid having a nucleotide sequence coding for *Dac g* I, or the functional equivalent of said nucleotide sequence.

61. An isolated nucleic acid sequence of claim 60 wherein said nucleotide sequence comprises the nucleotide sequence of Fig. 5.

62. An expression vector comprising a nucleotide sequence coding for *Dac g* I, or the functional equivalent of said nucleotide sequence.

63. A host cell transformed to express a protein encoded by the nucleic acid of claim 60.

64. Isolated *Dac g* I protein produced in a host cell transformed with the nucleic acid of claim 60.

65. An isolated nucleic acid having a nucleotide sequence coding for *Poa p* I, or the functional equivalent of said nucleotide sequence.

66. An isolated nucleic acid sequence of claim 65 wherein said nucleotide sequence comprises the nucleotide sequence of Fig. 6.

67. An expression vector comprising a nucleotide sequence coding for *Poa p* I, or the functional equivalent of said nucleotide sequence.

68. A host cell transformed to express a protein encoded by the nucleic acid of claim 65.

69. Isolated *Poa p* I protein produced in a host cell transformed with the nucleic acid of claim 60.

70. An isolated protein allergen that is immunologically related to *Lol p* I.

71. An isolated protein allergen of claim 70 wherein said protein allergen is *Dac g* I or *Poa p* I.

T CELL EPITOPES OF RYEGRASS POLLEN ALLERGEN

Background of the Invention

The most abundant proteins of grass pollen are allergens, which are the major cause of allergic disease in temperate climates (Marsh (1975), "Allergens and the genetics of allergy"; in M. Sela (ed), *The Antigens*, 3:271-359, Academic Press Inc., London, New York)., Hill *et al.* (1979) *Medical Journal of Australia,* 1:426-429). The first descriptions of the allergenic proteins in ryegrass showed that they are immunochemically distinct, and are known as groups I, II, III and IV (Johnson and Marsh (1965), *Nature*, 206:935-942; and Johnson and Marsh (1966) *Immunochemistry,* 3:91-100). Using the International Union of Immunological Societies' (IUIS) nomenclature, these allergens are designated *Lol p* I, *Lol p* II, *Lol p* III, and *Lol p* IV. Another important *Lolium perenne* allergen which has been identified in the literature is *Lol p* IX, also known as *Lol p* V or *Lol p* Ib, which has been found to be closely related to the Group V protein allergens in grasses.

These proteins have been identified in pollen from ryegrass, *Lolium perenne*, and act as antigens in triggering immediate (Type 1) hypersensitivity in susceptible humans.

~~*Lol p* I is defined as an allergen because of its ability to bind to specific IgE in~~ sera of ryegrass-sensitive patients, to act as an antigen in IgG responses and to trigger T-cell responses. The allergenic properties have been assessed by direct skin testing of grass pollen-sensitive patients. The results showed that 84% had a skin sensitivity to *Lol p* I (Freidhoff, et al., (1986) *J. Allergy Clin. Immunol.*, 78:1190-1201) demonstrating the primary importance of this protein as the major allergen. Furthermore, 95% of patients demonstrated to be grass pollen-sensitive possessed specific IgE antibody that bound to *Lol p* I, as demonstrated by immunoblotting (Ford and Baldo (1986) *International Archives of Allergy and Applied Immunology*, 81:193-203).

Substantial allergenic cross-reactivity between grass pollens has been demonstrated using an IgE-binding assay, the radioallergo-sorbent test (RAST), for example, as described by Marsh *et al.* (1970) *J. Allergy*, 46:107-121, and Lowenstein (1978) *Prog. Allergy*, 25:1-62. (Karger, Basel).

The immunochemical relationship of *Lol p* I with other grass pollen antigens has been demonstrated using both polyclonal and monoclonal antibodies (e.g., Smart and Knox (1979) *International Archives of Allergy and*

Applied Immunology, 62: 173-187; Singh and Knox (1985), *International Archives of Allergy and Applied Immunology*, 78:300-304). Antibodies have been prepared to both purified proteins and IgE-binding components. These data demonstrate that the major allergen present in pollen of closely related grasses is immunochemically similar to *Lol p* I (Singh and Knox, *supra*). Grasses that may be considered immunochemically related to *Lol p* I and that comprise allergens which may be considered immunologically cross-reactive with antibody to *Lol p* I include:

Pooid (festucoid) grasses of the Poaceae (Gramineae) family include the following. GROUP 1: Triticanea: *Bromus inermis*, smooth brome; *Agropyron repens*, English couch; *A. cristatum*; *Secale cereale* rye *Triticum aestivum*, wheat. GROUP 2: Poanae: *Dactylis glomerata*, orchard grass of cocksfoot; *Festuca elatior*, meadow fescue; *Lolium perenne*, perennial ryegrass; *L.multiflorum*, Italian ryegrass; *Poa pratensis*, Kentucky bluegrass; *P.compressa*, flattened meadow grass; *Avena sativa*, oat; *Holcus lanatus*, velvet grass or Yorkshire fog; *Anthoxanthum odoratum*; sweet vernal grass; *Arrhenatherum elatius*, oat grass; *Agrostis alba*, red top; *Phleum pratense*, timothy; *Phalaris arundinacea*, reed canary grass. *Panicoid* grass, *Paspalum notatum*, Bahia grass, Andropogonoid grasses: *Sorghum halepensis*, Johnson grass.

In view of the prevalence of ryegrass pollen allergens and related grass allergens all over the world, there is a pressing need for the development of compositions and methods that could be used in detecting sensitivities to *Lol p* I or other immunologically related grass allergens, or in treating sensitivities to such allergens, or in assisting in the manufacture of medicaments to treat such sensitivities. The present invention provides materials and methods having one or more of those utilities.

Summary of the Invention

The present invention provides isolated peptides of *Lol p* I. Peptides within the scope of the invention comprise at least one T cell epitope, preferably at least two T cell epitopes of *Lol p* I. The invention further provides peptides comprising at least two regions, each region comprising at least one T cell epitope of *Lol p* I.

The invention also provides modified peptides having similar or enhanced therapeutic or diagnostic properties as the corresponding, naturally-occurring allergen or portion thereof, but also having advantageous physical or biological properties, such as reduced side effects, reduced IgE binding, improved solubility, increased *in vitro* or *in vivo* T cell stimulating ability, increased stability or the like. Preferred peptides of the invention are capable of modifying, in a *Lol p* I-sensitive individual to whom they are administered, the allergic response of the individual to *Lol p* I or an allergen immunologically cross-reactive with *Lol p* I, e.g., allergens derived from pollen belonging to the Poaceae (Gramineae) family, such as *Dactylis glomerata* (*Dac g* I), *Poa pretensis* (*Poa p* I) and *Phleum pratense* (*Phl p* I), as discussed above.

The present invention also provides non-native (i.e., recombinant or chemically synthesized) *Lol p* I peptides or their derivatives or homologues and provides non-native allergenic protein or peptides immunologically cross-reactive with antibodies or with T cells of *Lol p* I or derivatives or homologues thereof.

The present invention also provides *Dac g* I and *Poa p* I protein allergens which are immunologically cross-reactive with *Lol p* I, and fragments of *Dac g* I and *Poa p* I produced in a host cell transformed with a nucleic acid sequence coding for *Dac g* I and *Poa p* I, respectively, and fragments of *Dac g* I and *Poa p* I prepared synthetically. The present invention further provides nucleic acid sequences coding for *Dac g* I, *Poa p* I and fragments thereof. Also provided are isolated peptides of *Dac g* I and *Poa p* I comprising at least one T cell epitope which are immunologically cross-reactive with peptides comprising at least one T cell epitope derived from *Lol p* I.

Methods of treatment and of diagnosis of sensitivity to ryegrass pollen protein, *Lol p* I, or to pollen proteins that are immunologically related to *Lol p* I (such as *Dac g* I, *Phl p* I and *Poa p* I), as well as compositions comprising one or more peptides of the invention, are also provided.

Further features of the present invention will be better understood from the following detailed description of the preferred embodiments of the invention in conjunction with the appended figures.

Brief Description of the Figures

Fig. 1 shows the nucleotide sequence of cDNA clone 26.*j* (SEQ ID NO 1) and its predicted amino acid sequence (SEQ ID NO: 2). Clone 26.*j* is a PCR-generated, full-length clone of *Lol p* I.

Fig. 2 shows various peptides of desired lengths derived from *Lol p* I (SEQ ID NO: 3-30); such peptides include polymorphisms inherent in the *Lol p* I sequence (i.e., LPI-4.1 (SEQ ID NO: 8) and LPI-16.1 (SEQ ID NO: 23)) or homologues of peptides derived from *Lol p* I (i.e., LPI-11 (SEQ ID NO: 15), and LPI-12 (SEQ ID NO: 17)).

Fig. 3 is a graphic representation depicting responses of T cell lines from thirty-five grass-sensitive patients primed *in vitro* with purified native *Lol p* I and analyzed for response to various *Lol p* I peptides by percent of positive responses (with an S.I. of at least two, shown over each bar), the mean stimulation index of positive response for the peptide (shown over each bar in parentheses) and the positivity index (% positive x mean S.I. index, Y axis).

Fig. 4 shows various peptides of desired lengths derived from *Lol p* I (SEQ ID NO: 23, 25, 27, 30-50).

Fig. 5 shows the nucleotide sequence of cDNA clone 106.5 (SEQ ID NO: 51) and its predicted amino acid sequence (SEQ ID NO: 52). Clone 106.5 is a PCR-generated, full-length clone of *Dac g* I.

Fig. 6 shows the nucleotide sequence of cDNA clone 114 (SEQ ID NO: 53) and its predicted amino acid sequence (SEQ ID NO: 54). Clone 114 is a PCR-generated, full-length clone of *Poa p* I.

Fig. 7 shows the nucleotide sequence of cDNA clone 20 (SEQ ID NO: 55) and its predicted amino acid sequence (SEQ ID NO: 56). Clone 20 is a PCR generated, full length clone of *Phl p* I.

Fig. 8 shows a comparison of the amino acid sequences of the mature protein of *Lol p* I (SEQ ID NO: 57), *Dac g* I (SEQ ID NO: 58), *Phl p* I (SEQ ID NO: 59), and *Poa p* I (SEQ ID NO: 60), including polymorphisms thereof.

Fig. 9 shows a comparison of various peptides comprising at least one T cell epitope derived from *Lol p* I, with homologous peptides derived from the same regions of *Dac g* I, *Phl p* I, and *Poa p* I (SEQ ID NO: 23, 25, 27, 30, 61-70).

Detailed Description of the Invention

The present invention provides isolated peptides derived from *Lol p* I (SEQ ID NO: 3-50). The present invention also provides *Dac g* I and *Poa p* I protein allergens which are immunologically cross-reactive with *Lol p* I. The term "peptide"

as used herein refers to any protein fragment of *Lol p* I that induces an immune response. The terms "fragment" and "antigenic fragment" of a protein as used interchangeably herein refer to an amino acid sequence having fewer amino acid residues than the entire native amino acid sequence of the protein from which the fragment is derived, and that induces an immune response. The terms "isolated" and "purified" as used herein refer to peptides of the invention which are substantially free of cellular material or culture medium when produced by recombinant DNA techniques, or substantially free of chemical precursors or other chemicals when synthesized chemically. Preferred peptides of the invention include peptides derived from *Lol p* I which comprise at least one T cell epitope of the allergen, or a portion of such a peptide which includes at least one T cell epitope.

Peptides comprising at least two regions, each region comprising at least one T cell epitope *Lol p* I are also within the scope of the invention. Isolated peptides or regions of isolated peptides, each comprising at least two T cell epitopes of the *Lol p* I protein allergen are particularly desirable for increased therapeutic effectiveness. Peptides that are immunologically related (e.g., by antibody or T cell cross-reactivity) to peptides of the present invention, such as peptides derived from *Dac g* I and *Poa p* I, are also within the scope of the invention. Peptides immunologically related by antibody cross-reactivity are recognized by antibodies specific for a peptide of *Lol p* I. Peptides immunologically related to a given peptide by T cell cross-reactivity are capable of also reacting with the same T cells that react with that given peptide.

Isolated protein and peptides of the invention can be produced by recombinant DNA techniques in a host cell transformed with a nucleic acid having a sequence encoding such peptide. The isolated peptides of the invention can also be produced by chemical synthesis. When a protein or peptide is produced by recombinant techniques, host cells transformed with a nucleic acid having a sequence encoding a peptide of the invention or the functional equivalent of the nucleic acid sequence are cultured in a medium suitable for the cells. Peptides can be purified from cell culture medium, host cells, or both, using techniques known in the art for purifying peptides and proteins including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis or immunopurification with antibodies specific for the peptide, the protein allergen from which the peptide is derived, or a portion thereof.

The present invention provides expression vectors and host cells transformed to express the nucleic acid sequences of the invention. Nucleic acids coding for *Lol p* I peptides of the invention, or at least a portion thereof, may be expressed in bacterial

cells such as *E. coli*, insect cells, yeast, or mammalian cells such as Chinese hamster ovary cells (CHO). Suitable expression vectors, promoters, enhancers, and other expression control elements may be found in Sambrook et al. *Molecular Cloning: A Laboratory Manual*, second edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989. Other suitable expression vectors, promoters, enhancers, and other expression elements are known to those skilled in the art. Expression in mammalian, yeast or insect cells leads to partial or complete glycosylation of the recombinant material and formation of any inter- or intra-chain disulfide bonds. Suitable vectors for expression in yeast include YepSec1 (Baldari et al. (1987) *Embo J.*, 6: 229-234); pMFa (Kurjan and Herskowitz (1982) *Cell*, 30: 933-943); JRY88 (Schultz et al. (1987) *Gene*, 54: 113-123) and pYES2 (Invitrogen Corporation, San Diego, CA). These vectors are freely available. Baculovirus and mammalian expression systems are also available. For example, a baculovirus system is commercially available (PharMingen, San Diego, CA) for expression in insect cells while the pMSG vector is commercially available (Pharmacia, Piscataway, NJ) for expression in mammalian cells.

For expression in *E. coli*, suitable expression vectors include, among others, pTRC (Amann et al. (1988) *Gene*, 69: 301-315); pGEX (Amrad Corp., Melbourne, Australia); pMAL (N.E. Biolabs, Beverly, MA); pRIT5 (Pharmacia, Piscataway, NJ); pET-11d (Novagen, Madison, WI) Jameel et al., (1990) *J. Virol.*, 64:3963-3966; and pSEM (Knapp et al. (1990) *BioTechniques*, 8: 280-281). The use of pTRC, and pET-11d, for example, will lead to the expression of unfused protein. The use of pMAL, pRIT5 pSEM and pGEX will lead to the expression of allergen fused to maltose E binding protein (pMAL), protein A (pRIT5), truncated ß-galactosidase (PSEM), or glutathione S-transferase (pGEX). When a *Lol p* I peptide of the invention, is expressed as a fusion protein, it is particularly advantageous to introduce an enzymatic cleavage site at the fusion junction between the carrier protein and the *Lol p* I peptide. The *Lol p* I peptide may then be recovered from the fusion protein through enzymatic cleavage at the enzymatic site and biochemical purification using conventional techniques for purification of proteins and peptides. Suitable enzymatic cleavage sites include those for blood clotting Factor Xa or thrombin for which the appropriate enzymes and protocols for cleavage are commercially available from, for example, Sigma Chemical Company, St. Louis, MO and N.E. Biolabs, Beverly, MA. The different vectors also have different promoter regions allowing constitutive or inducible expression with, for example, IPTG induction (PRTC, Amann et al., (1988)

supra; pET-11d, Novagen, Madison, WI) or temperature induction (pRIT5, Pharmacia, Piscataway, NJ) . It may also be appropriate to express recombinant *Lol p* I peptides in different *E. coli* hosts that have an altered capacity to degrade recombinantly expressed proteins (e.g., U.S. Patent 4,758,512). Alternatively, it may be advantageous to alter the nucleic acid sequence to use codons preferentially utilized by *E. coli*, where such nucleic acid alteration would not affect the amino acid sequence of the expressed protein.

Host cells can be transformed to express the nucleic acid sequences of the invention using conventional techniques such as calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, or electroporation. Suitable methods for transforming the host cells may be found in Sambrook et al. *supra*, and other laboratory textbooks. The nucleic acid sequences of the invention may also be chemically synthesized using standard techniques (i.e., solid phase synthesis). Details of the cloning of *Lol p* I are given in the Examples.

Inducible non-fusion expression vectors include pTrc (Amann *et al.*, (1988) *Gene*, **69**:301-315) and pET11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology*, Academic Press, San Diego, California (1990), **185**:60-89). While target gene expression relies on host RNA polymerase transcription from the hybrid trp-lac fusion promoter in pTrc, expression of target genes inserted into pET11d relies on transcription from the T7 gn10-lac 0 fusion promoter mediated by coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident λ prophage harboring a T7 gn1 under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant *Lol p* I peptide expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology*, Academic Press, San Diego, California (1990), **185**:119-128). Another strategy would be to alter the nucleic acid sequence of the desired gene to be inserted into an expression vector so that the individual codons for each amino acid would be those preferentially utilized in highly expressed *E. coli* proteins (Wada *et al.*, (1992) *Nuc. Acids Res.* **20**:2111-2118). Such alteration of nucleic acid sequences of the invention could be carried out by standard DNA synthesis techniques.

The nucleic acids of the invention can also be chemically synthesized using standard techniques. Various methods of chemically synthesizing

polydeoxynucleotides are known, including solid-phase synthesis which, like peptide synthesis, has been fully automated in commercially available DNA synthesizers (See e.g., Itakura *et al.* U.S. Patent 4,598,049; Caruthers *et al.* U.S. Patent 4,458,066; and Itakura U.S. Patents 4,401,796 and 4,373,071, incorporated by reference herein).

The present invention also provides fragments of nucleic acid sequences encoding peptides of the invention. As used herein, the term "fragment" of a nucleic acid sequence refers to a nucleotide sequence having fewer bases than the nucleotide sequence coding for the entire amino acid sequence of the protein. Nucleic acid sequences used in any embodiment of this invention can be cDNA obtained as described herein, or alternatively, can be any oligodeoxynucleotide sequence having all or a portion of a sequence represented herein, or their functional equivalents. Such oligodeoxynucleotide sequences can be produced chemically or mechanically, using known techniques. A functional equivalent of an oligonucleotide sequence of *Lol p* I is one which is 1) a sequence capable of hybridizing to a complementary oligonucleotide to which the sequence (or corresponding sequence portions) of *Lol p* I as shown in Fig. 1 (SEQ ID NO: 1) or fragments thereof hybridizes, or 2) the sequence (or corresponding sequence portion) complementary to the sequence of *Lol p* I as shown in Fig. 1 (SEQ ID NO: 1), and/or 3) a sequence which encodes a product (e.g., a polypeptide or peptide) having the same functional characteristics of the product encoded by the sequence (or corresponding sequence portion) of *Lol p* I as shown in Fig. 1 (SEQ ID NO: 1). Whether a functional equivalent must meet one or both criteria will depend on its use (e.g., if it is to be used only as an oligonucleotide probe, it need meet only the first or second criteria and if it is to be used to produce a *Lol p* I peptide of the invention, it need only meet the third criterion).

Preferred nucleic acids encode a peptide having at least about 50% homology to a *Lol p* I peptide of the invention, more preferably at least about 60% homology and most preferably at least about 70% homology with a *Lol p* I peptide of the invention. Nucleic acids that encode peptides having at least about 90%, more preferably at least about 95%, and most preferably at least about 98-99% homology with *Lol p* I peptides of the invention are also within the scope of the invention. Homology refers to sequence similarity between two peptides of *Lol p* I, or between two nucleic acid molecules. Homology can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same nucleotide or amino acid, then molecules are

homologous at that position. A degree of homology between sequences is a function of the number of matching or homologous positions shared by the sequences.

Preferred nucleic acid fragments encode peptides of at least 7 amino acid residues in length, and preferably 13-40 amino acid residues in length, and more preferably at least 16-30 amino acids residues in length, Nucleic acid fragments encoding peptides of at least 30 amino acid residues in length, at least 40 amino acid residues in length, at least about 80 amino acid residues in length, at least about 100 amino acid residues in length or more, are also contemplated.

Also within the scope of the invention are nucleic acid sequences encoding allergens immunologically cross-reactive with *Lol p* I, such as full length *Dac g* I and *Poa p* I proteins or peptides (Figs 5 (SEQ ID NO: 52) , 6 (SEQ ID NO: 54) , and 9 (SEQ ID NO: 23, 25, 27, 30, 61-70)). Proteins and peptides of *Dac g* I and *Poa p* I may be produced recombinantly as discussed above, or synthetically. Expression vectors and host cells transformed to express *Dac g* I and *Poa p* I proteins or peptides thereof are also within the scope of the invention. Details of the cloning of *Dac g* I and *Poa p* I are given in the examples.

The present invention also provides a method of producing isolated *Lol p* I peptides of the invention or a portion thereof, comprising the steps of culturing a host cell transformed with a nucleic acid sequence encoding a *Lol p* I peptide of the invention in an appropriate medium to produce a mixture of cells and medium containing said *Lol p* I peptide; and purifying the mixture to produce substantially pure *Lol p* I peptide. Host cells transformed with an expression vector containing DNA coding for a *Lol p* I peptide of the invention are cultured in a suitable medium for the host cell. *Lol p* I peptides of the invention can be purified from cell culture medium, host cells, or both using techniques known in the art for purifying peptides and proteins including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis and immunopurification with antibodies specific for the *Lol p* I peptides or portions thereof.

Another aspect of the present invention pertains to an antibody specifically reactive with a *Lol p* I peptide. Such antibodies may be used to standardize allergen extracts or to isolate the naturally occurring *Lol p* I . Also, Lol p I peptides of the invention can be used as "purified" allergens to standardize allergen extracts. For example, an animal such as a mouse or rabbit can be immunized with an immunogenic form of an isolated *Lol p* I peptide of the invention capable of eliciting an antibody response. Techniques for conferring immunogenicity on a peptide include conjugation

to carriers or other techniques well-known in the art. The *Lol p* I peptide can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum standard ELISA or other immunoassay can be used with the immunogen as antigen to assess the levels of antibodies.

Following immunization, anti-*Lol p* I peptide antisera can be obtained and, if desired, polyclonal anti-*Lol p* I peptide antibodies from the serum. To produce monoclonal antibodies, antibody producing cells (lymphocytes) can be harvested from an immunized animal and fused by standard somatic cell fusion procedures with immortalizing cells such as myeloma cells to yield hybridoma cells. Hybridoma cells can be screened immunochemically for production of antibodies reactive with the *Lol p* I peptides of the invention. These sera or monoclonal antibodies can be used to standardize allergen extracts.

Through use of the peptides and antibodies of the present invention, preparations of consistent, well-defined composition and uniform biological activity can be made. Compositions having therapeutic activity may be administered for therapeutic purposes (e.g., to modify the allergic response of a ryegrass sensitive individual to pollen of such grasses or pollen of an immunologically related grass such as ~~*Dac g* I, *Poa p* I and *Phl p* I~~). Administration of such peptides may, for example, modify B-cell response to *Lol p* I allergen, T-cell response to *Lol p* I allergen or both responses. Isolated peptides can also be used to study the mechanism of immunotherapy of ryegrass pollen allergy and to design modified derivatives or analogues useful in immunotherapy. Compositions according to the invention will have utility in diagnosis of ryegrass sensitivity or sensitivity to grass allergens cross-reactive to ryegrass allergens, because the components include T cell epitopes recognizing the allergens.

The present invention also pertains to T cell clones which specifically recognize *Lol p* I peptides of the invention. These T cell clones may be suitable for isolation and molecular cloning of the gene for the T cell receptor which is specifically reactive with a peptide of the present invention. The T cell clones may be produced as described in Example 4, or as described in *Cellular Molecular Immunology,* Abdul K. Abbas et al., W.B. Saunders Co. (1991) pg. 139. The present invention also pertains to soluble T cell receptors. These receptors may inhibit antigen-dependent activation of the relevant T cell subpopulation within an individual sensitive to *Lol p* I. Antibodies specifically reactive with such a T cell receptor can also be produced according to the

techniques described herein. Such antibodies may also be useful to block T-cell-MHC interaction in an individual. Methods for producing soluble T cell receptors are described in *Immunology: A Synthesis,* 2nd Ed., Edward S. Golub *et al.*, Sinaur Assoc., Sunderland, Massachusetts, (1991) pp. 366-369.

It is also possible to modify the structure of a peptide of the invention to achieve additional advantageous physical or biological properties such as increasing solubility, enhancing therapeutic or preventive efficacy, increasing stability (e.g., shelf life *ex vivo* or resistance to proteolytic degradation *in vivo*), decreasing adverse side effects, and the like. A modified peptide can be produced in which the amino acid sequence has been altered, such as by amino acid substitution, deletion, or addition, in order to modify immunogenicity and/or to reduce allergenicity. Peptides may also be advantageously modified by addition or conjugation with another peptide or other component.

For example, a peptide can be modified so that it maintains the ability to induce T cell anergy and to bind MHC proteins but reduces the ability to induce a strong proliferative response, or possibly any proliferative response, when administered in immunogenic form. In this instance, critical binding residues for the T cell receptor can be determined using known techniques (e.g., substitution of each residue and determination of the presence or absence of T cell reactivity). Those residues shown to be essential to interact with the T cell receptor can be modified by replacing the essential amino acid with another preferably similar amino acid residue (a "conservative substitution") whose presence is shown to enhance, diminish but not eliminate, or not affect T cell reactivity. In addition, those amino acid residues that are not essential for T cell receptor interaction can be modified by replacement with another amino acid whose incorporation may enhance, diminish or not affect T cell reactivity but does not eliminate binding to relevant MHC.

Additionally, peptides of the invention can be modified by replacing an amino acid shown to be essential to interact with the MHC protein complex with another, preferably similar amino acid residue (conservative substitution) whose presence is shown to enhance, diminish but not eliminate or not affect T cell reactivity. In addition, amino acid residues that are not essential for interaction with the MHC protein complex but that still bind the MHC protein complex can be modified by replacement with another amino acid whose incorporation may enhance, not affect, or diminish but not eliminate T cell reactivity. Preferred amino acid substitutions for non-

essential amino acids include, but are not limited to substitutions with alanine, glutamic acid, or a methyl amino acid.

In order to enhance stability and/or reactivity, peptides of the invention can also be modified to incorporate one or more polymorphisms in the amino acid sequence of the protein allergen resulting from natural allelic variation. Additionally, D-amino acids, non-natural amino acids or non-amino acid analogues can be substituted or added to produce a modified peptide within the scope of this invention. Furthermore, peptides of the present invention can be modified using the polyethylene glycol (PEG) method of A. Sehon and co-workers (Wie et al., *supra*) to produce a protein or peptide conjugated with PEG. In addition, PEG can be added during chemical synthesis of a protein or peptide of the invention. Modifications of peptides or portions thereof can also include reduction/ alyklation (Tarr in: *Methods of Protein Microcharacterization,* J.E. Silver ed. Humana Press, Clifton, NJ, pp 155-194 (1986)); acylation (Tarr, *supra*); chemical coupling to an appropriate carrier (Mishell and Shiigi, eds, *Selected Methods in Cellular Immunology,* WH Freeman, San Francisco, CA (1980); U.S. Patent 4,939,239; or mild formalin treatment (Marsh *International Archives of Allergy and Applied Immunology,* **41**:199-215 (1971)).

To facilitate purification and potentially increase solubility of peptides of the invention, it is possible to add reporter group(s) to the peptide backbone. For example, poly-histidine can be added to a peptide to purify the peptide by immobilized metal ion affinity chromatography (Hochuli, E. et al., *Bio/Technology,* 6:1321-1325 (1988)). In addition, specific endoprotease cleavage sites can be introduced, if desired, between a reporter group and amino acid sequences of a peptide to facilitate isolation of peptides free of irrelevant sequences. In order to successfully desensitize an individual to a protein antigen, it may be necessary to increase the solubility of a peptide by adding functional groups to the peptide or by not including hydrophobic T cell epitopes or regions containing hydrophobic epitopes in the peptides or hydrophobic regions of the protein or peptide. Functional groups such as charged amino acid pairs (e.g., KK or RR) are particularly useful for increasing the solubility of a peptide when added to the amino or carboxy terminus of the peptide. Examples of modifications to peptides to increase solubility include modifications to peptide LPI-16.1 (SEQ ID NO: 23) (Fig. 2), such modified peptides include: LPI-16.2 (SEQ ID NO: 31), LP1-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID NO 33) , LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID NO: 35), LPI-16.7 (SEQ ID NO: 36) , LPI-16.9 (SEQ ID NO: 37), LPI-16.10 (SEQ ID NO: 38), all as shown in Fig. 4.

To potentially aid proper antigen processing of T cell epitopes within a peptide, canonical protease sensitive sites can be recombinantly or synthetically engineered between regions, each comprising at least one T cell epitope. For example, charged amino acid pairs, such as KK or RR, can be introduced between regions within a peptide during recombinant construction of the peptide or added to the amino or carboxy terminus of a synthetically produced peptide. The resulting peptide can be rendered sensitive to cathepsin and/or other trypsin-like enzymes cleavage to generate portions of the peptide containing one or more T cell epitopes. In addition, as mentioned above, such charged amino acid residues can result in an increase in solubility of a peptide.

Site-directed mutagenesis of DNA encoding a peptide of the invention can be used to modify the structure of the peptide by methods known in the art. Such methods may, among others, include PCR with degenerate oligonucleotides (Ho et al., *Gene,* 77:51-59 (1989)) or total synthesis of mutated genes (Hostomsky, Z. et al., *Biochem. Biophys, Res. Comm.*, **161**:1056-1063 (1989)). To enhance bacterial expression, the aforementioned methods can be used in conjunction with other procedures to change the eucaryotic codons in DNA constructs encoding protein or peptides of the invention to ones preferentially processed in *E. coli*, yeast, mammalian cells, or other prokaryotic or eukaryotic host cells.

Peptides of the present invention can also be used for detecting and diagnosing ryegrass pollinosis. For example, this could be done *in vitro* by combining blood or blood products obtained from an individual to be assessed for sensitivity to ryegrass pollen or another cross-reactive pollen such as *Dac g* I, *Poa p* I and *Phl p* I, with an isolated peptide(s) of *Lol p* I, under conditions appropriate for binding of components in the blood (e.g., antibodies, T-cells, B cells) with the peptide(s) and determining the extent to which such binding occurs. Other diagnostic methods for allergic diseases in which the protein, peptides or antibodies of the present invention will be useful include radio-allergergosorbent test (RAST), paper radioimmunosorbent test (PRIST), enzyme linked immunosorbent assay (ELISA), radioimmunoassays (RIA), immuno-radiometric assays (IRMA), luminescence immunoassays (LIA), histamine release assays and IgE immunoblots.

The presence in individuals of IgE specific for at least one protein allergen and the ability of T cells of the individuals to respond to T cell epitope(s) of the protein allergen can be determined by administering to the individuals an Immediate Type Hypersensitivity test and a Delayed Type Hypersensitiity test. The individuals

are administered an Immediate Type Hypersensitivity test (see e.g., *Immunology* (1985) Roitt, I.M., Brostoff, J., Male, D.K. (eds), C.V. Mosby Co., Gower Medical Publishing, London, NY, pp. 19.2-19.18; pp. 22.1-22.10) utilizing the protein allergen or a portion thereof, or a modified form of the protein allergen or a portion thereof, each of which binds IgE specific for the allergen. The same individuals are administered a Delayed Type Hypersensitivity test prior to, simultaneously with, or subsequent to administration of the Immediate Type Hypersensitivity test. Of course, if the Immediate Type Hypersensitivity test is administered prior to the Delayed Type Hypersensitivity test, the Delayed Type Hypersensitivity test would be given to those individuals exhibiting a specific Immediate Type Hypersensitivity reaction. The Delayed Type Hypersensitivity test utilizes a modified form of the protein allergen or a portion thereof, the protein allergen produced recombinantly, or a peptide derived from the protein allergen, each of which has human T cell stimulating activity and each of which does not bind IgE specific for the allergen in a substantial percentage of the population of individuals sensitive to the allergen (e.g., at least about 75%). Those individuals found to have both a specific Immediate Type Hypersensitivity reaction and a specific Delayed Type Hypersensitivity reaction may be treated with a therapeutic composition comprising the same modified form of the protein or portion thereof, the recombinantly produced protein allergen, or the peptide, each as used in the Delayed Type Hypersensitivity test.

Isolated peptides of the invention, when administered in a therapeutic regimen to a *Lol p* I-sensitive individual (or an individual allergic to an allergen cross-reactive with ryegrass pollen allergen such as *Dac g* I, *Poa p* I and *Phl p* I) are capable of modifying the allergic response of the individual to *Lol p* I ryegrass pollen allergen (or such cross-reactive allergen). Preferably peptides of this invention are capable of modifying the B-cell response, T-cell response or both the B-cell and the T-cell response of the individual to the allergen. As used herein, modification of the allergic response of an individual sensitive to a ryegrass pollen allergen or cross-reactive allergen can be defined as non-responsiveness or diminution in symptoms to the allergen, as determined by standard clinical procedures (See, e.g., Varney et al, *British Medical Journal*, 302:265-269 (1990)) including diminution in ryegrass pollen-induced asthmatic symptoms. As referred to herein, a diminution in symptoms includes any reduction in allergic response of an individual to the allergen after the individual has completed a treatment regimen with a peptide or protein of the invention. This diminution may be subjective (i.e., the patient feels more comfortable

in the presence of the allergen), or diminution in symptoms may be determined clinically, using standard skin tests known in the art and discussed above.

Lol p I peptides of the present invention having T cell stimulating activity, and thus comprising at least one T cell epitope, are particularly preferred. In referring to an epitope, the epitope will be the basic element or smallest unit of recognition by a receptor, particularly immunoglobulins, histocompatibility antigens and T cell receptors where the epitope comprises amino acids essential to receptor recognition. Amino acid sequences which mimic those of the epitopes and which are capable of down-regulating or reducing allergic response to *Lol p* I can also be used. T cell epitopes are believed to be involved in initiation and perpetuation of the immune response to a protein allergen that is responsible for the clinical symptoms of allergy. Such T cell epitopes are thought to trigger early events at the level of the T helper cell by binding to an appropriate HLA molecule on the surface of an antigen presenting cell and stimulating the relevant T cell subpopulation. These events lead to T cell proliferation, lymphokine secretion, local inflammatory reactions, recruitment of additional immune cells to the site, and activation of the B cell cascade leading to production of antibodies. One isotype of these antibodies, IgE, is fundamentally important to the development of allergic symptoms, and its production is influenced early in the cascade of events, at the level of the T helper cell, by the nature of the lymphokines secreted.

Exposure of ryegrass pollen-sensitive patients or patients sensitive to an immunogically cross-reactive protein allergen such as *Dac g* I, *Poa p* I and *Phl p* I, to isolated *Lol p* I peptides of the present invention which comprise at least one T cell epitope and are derived from *Lol p* I protein allergen, may tolerize or anergize appropriate T cell subpopulations such that they become unresponsive to the protein allergen and do not participate in stimulating an immune response upon such exposure. In addition, administration of a peptide of the invention or portion thereof which comprises at least one T cell epitope may modify the lymphokine secretion profile as compared with exposure to the naturally-occurring *Lol p* I protein allergen or portion thereof (e.g., may result in a decrease of IL-4 and/or an increase in IL-2). Furthermore, exposure to such peptide of the invention may influence T cell subpopulations which normally participate in the response to the naturally occurring allergen such that these T cells are drawn away from the site(s) of normal exposure to the allergen (e.g., nasal mucosa, skin, and lung) towards the site(s) of therapeutic administration of the fragment or protein allergen. This redistribution of T cell

subpopulations can have the effect of ameliorating or reducing the ability of an individual's immune system to stimulate the usual immune response at the site of normal exposure to the allergen, resulting in a dimunution in allergic symptoms.

The isolated *Lol p* I peptides of the invention can be used in methods of diagnosing, treating or preventing allergic reactions to *Lol p* I allergen or an immunogically related protein allergen such as *Dac g* I, *Poa p* I and *Phl p* I. Thus, the present invention provides compositions useful in allergery diagnosis and/or useful in allergy therapy comprising isolated *Lol p* I peptides or portions thereof. Such compositions will typically also comprise a pharmaceutically acceptable carrier or diluent when intended for *in vivo* administration. Therapeutic compositions of the invention may include synthetically prepared *Lol p* I peptides.

Administration of the therapeutic compositions of the present invention to an individual to be desensitized can be carried out using known techniques. *Lol p* I peptides or portions thereof may be administered to an individual in combination with, for example, an appropriate diluent, a carrier and/or an adjuvant. Pharmaceutically acceptable diluents include saline and aqueous buffer solutions. Pharmaceutically acceptable carriers include polyethylene glycol (Wie *et al.* (1981) *Int. Arch. Allergy Appl. Immunol.*, 64:84-99) and liposomes (Strejan *et al.* (1984) *J. Neuroimmunol.*, 7:27). For purposes of inducing T cell anergy, the therapeutic composition is preferably administered in nonimmunogenic form, i.e., it does not contain adjuvant. The therapeutic compositions of the invention are administered to ryegrass pollen-sensitive individuals or individuals sensitive to an allergen which is immunologically cross-reactive with ryegrass pollen allergen (i.e., *Dactylis glomerata*, or *Sorghum halepensis*, etc.). Therapeutic compositions of the invention may also be used in the manufacture of medicaments for treating sensitivity to ryegrass pollen allergen or an immunologically related pollen allergen.

Administration of the therapeutic compositions of the present invention to an individual to be desensitized can be carried out using known procedures at dosages and for periods of time effective to reduce sensitivity (i.e., to reduce the allergic response) of the individual to the allergen. Effective amounts of the therapeutic compositions will vary according to factors such as the degree of sensitivity of the individual to ryegrass pollen, the age, sex, and weight of the individual, and the ability of the protein or fragment thereof to elicit an antigenic response in the individual.

The active compound (i.e., protein or fragment thereof) may be administered in any convenient manner such as by injection (subcutaneous, intravenous, etc.), oral

administration, inhalation, transdermal application, or rectal administration. Depending on the route of administration, the active compound may be coated within a material to protect the compound from the action of enzymes, acids and other natural conditions which may inactivate the compound.

For example, preferably about 1 μg- 3 mg and more preferably from about 20-750 μg of active compound (i.e., protein or fragment thereof) per dosage unit may be administered by injection. Dosage regimen may be adjusted to provide the optimum therapeutic response. For example, several divided doses may be administered daily or the dose may be proportionally reduced as indicated by the exigencies of the therapeutic situation.

To administer a peptide by other than parenteral administration, it may be necessary to coat the protein with, or co-administer the protein with, a material to prevent its inactivation. For example, the peptide or portion thereof may be co-administered with enzyme inhibitors or in liposomes. Enzyme inhibitors include pancreatic trypsin inhibitor, diisopropylfluorophosphate (DEP) and trasylol. Liposomes include water-in-oil-in-water CGF emulsions as well as conventional liposomes (Strejan *et al.*, (1984), *J. Neuroimmunol.*, 7:27).

The active compound may also be administered parenterally or intraperitoneally. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations may contain a preservative to prevent the growth of microorganisms.

Pharmaceutical compositions suitable for injection include sterile aqueous solutions (where the peptides are water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. In all cases, the composition intended for *in vivo* use must be sterile and must be fluid to the extent necessary to provide easy syringability. It should preferably be stable under the conditions of manufacture and storage and be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion, and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal

agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol and sorbitol or sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about, including in the composition, an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (i.e., protein or peptide) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile indectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient (i.e., protein or peptide) plus any additional desired ingredient from a previously sterile-filtered solution thereof.

When a peptide of the invention is suitably protected, as described above, the peptide may be orally administered, for example, with an inert diluent or an assimilable edible carrier. The peptide and other ingredients may also be enclosed in a hard or soft gelatin capsule, compressed into tablets, or incorporated directly into the individual's food. For oral therapeutic administration, the active compound may be formulated with conventional excipients and used in the form of ingestible tablets, buccal tablets, troches, capsules, elixirs, suspensions, syrups, wafers, and the like. Such compositions and preparations should contain at least 1% by weight of active compound. The percentage of the composition and preparations may, of course, be varied and may conveniently be between about 5 to 80% by weight of the dosage unit. The amount of active compound in such therapeutically useful compositions is such that a suitable dosage will be obtained. Preferred compositions or preparations according to the present invention are prepared so that an oral dosage unit contains from about 10 μg to about 200 mg of active compound.

The tablets, troches, pills, capsules and the like may also contain the following: a binder such as gum gragacanth, acacia, corn starch or gelatin; excipients such as dicalcium phosphate; a disintegrating agent such as corn starch, potato starch, alginic acid and the like; a lubricant such as magnesium stearate; and a sweetening agent such as sucrose, lactose or saccharin or a flavoring agent such as peppermint, oil of

wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar or both. A syrup or elixir may contain the active compound, sucrose as a sweetening agent, methyl and propylparabens as preservative, a dye and flavoring such as cherry or orange flavor. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compound may be incorporated into sustained-release preparations and formulations.

As used herein "pharmaceutically acceptable carrier" includes any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the therapeutic compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

Various isolated peptides of the invention derived from ryegrass pollen protein *Lol p* I are shown in Figs. 2 and 4 (SEQ ID NO: 3-50). Peptides comprising at least two regions, each region comprising at least one T cell epitope of *Lol p* I are also within the scope of the invention. As used herein a region may include the amino acid sequence of a peptide of the invention as shown in Figs. 2 and 4 (SEQ ID NO: 3-50) or the amino acid sequence of a portion of such peptide.

To obtain isolated peptides of the present invention, *Lol p I* is divided into non-overlapping peptides of desired length or overlapping peptides of desired lengths as discussed in Example 4 which can be produced recombinantly, or synthetically. Peptides comprising at least one T cell epitope are capable of eliciting a T cell response, such as T cell proliferation or lymphokine secretion, and/or are capable of inducing T cell anergy (i.e., tolerization). To determine peptides comprising at least one T cell epitope, isolated peptides are tested by, for example, T cell biology techniques, to determine whether the peptides elicit a T cell response or induce T cell anergy. Those peptides found to elicit a T cell response or to induce T cell anergy are defined as having T cell stimulating activity.

As discussed in Example 4, human T cell stimulating activity can be tested by culturing T cells obtained from an individual sensitive to *Lol p* I allergen, (i.e., an

individual who has an IgE-mediated immune response to *Lol p* I allergen) with a peptide derived from the allergen, then determining whether proliferation of T cells occurs in response to the peptide. T cell proliferation may be measured in several ways, e.g., by cellular uptake of tritiated thymidine. Stimulation indices for responses by T cells to peptides can be calculated as the maximum counts-per-minute (CPM) in response to a peptide divided by the control CPM. A stimulation index (S.I.) equal to or greater than two times the background level is considered "positive". Positive results are used to calculate the mean stimulation index for each peptide for the group of patients tested. Preferred peptides of this invention comprise at least one T cell epitope and have a mean T cell stimulation index of greater than or equal to 2.0. A peptide having a mean T cell stimulation index of greater than or equal to 2.0 in a significant number of ryegrass pollen sensitive patients tested (i.e., at least 10% of patients tested) is considered useful as a therapeutic agent. Preferred peptides have a mean T cell stimulation index of at least 2.5, more preferably at least 3.0, more preferably at least 3.5, more preferably at least 4.0, more preferably at least 5, and most preferably at least about 6. For example, peptides of the invention having a mean T cell stimulation index of at least 5, as shown in Fig. 3, include LPI-2 (SEQ ID NO: 5), LPI-3 (SEQ ID NO: 6), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-17 (SEQ ID NO: 24), LPI-19 (SEQ ID NO: 26), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29) and LPI-23 (SEQ ID NO: 30). For example, peptides of the invention having a mean T cell stimulation index of at least 6, as shown in Fig. 3, include LPI-2 (SEQ ID NO: 5), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30).

In addition, preferred peptides have a positivity index (P.I.) of at least about 100, more preferably at least about 200 and most preferably at least about 300. The positivity index for a peptide is determined by multiplying the mean T cell stimulation index by the percent of individuals, in a population of individuals sensitive to ryegrass pollen (e.g., preferably at least 15 individuals, more preferably at least 30 individuals or more), who have a T cell stimulation index to such peptide of at least 2.0. Thus, the positivity index represents both the strength of a T cell response to a peptide (S.I.) and the frequency of a T cell response to a peptide in a population of individuals sensitive to ryegrass pollen. For example, as shown in Fig. 3, *Lol p* I peptide LPI-15 (SEQ ID NO: 21) has a mean S.I. of 12.2 and 11% of positive responses in the group of individuals tested resulting in a positivity index of 134.2. *Lol p* I peptides having a

positivity index of at least about 100 and a mean T cell stimulation index of at least about 4 include: LPI-2 (SEQ ID NO: 5), LPI-11 (SEQ ID NO: 15), LPI-13 (SEQ ID NO: 19), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30).

In order to determine precise T cell epitopes by, for example, fine mapping techniques, a peptide having T cell stimulating activity and thus comprising at least one T cell epitope as determined by T cell biology techniques is modified by addition or deletion of amino acid residues at either the amino or carboxy terminus of the peptide and tested to determine a change in T cell reactivity to the modified peptide. If two or more peptides which share an area of overlap in the native protein sequence are found to have human T cell stimulating activity, as determined by T cell biology techniques, additional peptides can be produced comprising all or a portion of such peptides and these additional peptides can be tested by a similar procedure. Following this technique, peptides are selected and produced recombinantly or synthetically. Examples of fine map peptides are as follows: modified versions of peptide LPI-18 (SEQ ID NO: 25) (Fig. 2) include peptides: LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-18.8 (SEQ ID NO: 42) all as shown in Fig. 4; modified versions of peptide LPI-20 (SEQ ID NO: 27) (Fig. 2) include peptides: LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), and LPI-20.6 (SEQ ID NO: 47) all as shown in Fig. 4; modified versions of peptide LPI-23 (SEQ ID NO: 30) (Fig. 2) include peptides: LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49) and LPI-23.4 (SEQ ID NO: 50) all as shown in Fig. 4.

Peptides are selected for diagnostic or therapeutic uses based on various factors, including the strength of the T cell response to the peptide (e.g., stimulation index), the frequency of the T cell response to the peptide in a population of individuals sensitive to ryegrass pollen, and the potential cross-reactivity of the peptide with other allergens from other species of grasses as discussed earlier. The physical and chemical properties of these selected peptides (e.g., solubility, stability) are examined to determine whether the peptides are suitable for use in therapeutic compositions or whether the peptides require modification as described herein. The ability of the selected peptides or selected modified peptides to stimulate human T cells (e.g., induce proliferation, lymphokine secretion) is determined.

The most preferred T cell epitope-containing peptides of the invention do not bind immunoglobulin E (IgE) of an allergic individual or bind IgE to a substantially lesser extent (e.g., at least 100 fold less and more preferably, at least 1000 fold less) than the protein allergen from which the peptide is derived. The major complications of standard immunotherapy are IgE-mediated responses such as anaphylaxis. Immunoglobulin E is a mediator of anaphylactic reactions which result from the binding and cross-linking of antigen to IgE on mast cells or basophils and the consequent release of mediators (e.g., histamine, serotonin, eosinophil chemotacic factors). Anaphylaxis in a substantial percentage of a population of individuals sensitive to *Lol p* I could be avoided by the use in immunotherapy of a peptide which do not bind IgE in a substantial percentage (e.g., at least about 75%) of a population of individuals sensitive to *Lol p* I allergen, or, if the peptides do bind IgE, such binding does not result in the release of mediators from mast cells or basophils. The risk of anaphylaxis could be reduced by the use in immunotherapy of a peptide or peptides which have reduced IgE binding. Moreover, peptides having minimal IgE stimulating activity are desirable for therapeutic effectiveness. Minimal IgE stimulating activity refers to IgE production that is less than the amount of IgE production stimulated by the native *Lol p* I protein allergen. Similarly, IL-4 production can be compared, with reduces IL-4 production indicating lessened IgE stimulating activity.

Preferred T cell epitope-containing peptides of the invention, when administered to a ryegrass pollen-sensitive individual or an individual sensitive to an allergen which is immunologically related to ryegrass pollen allergen (such as *Dac g* I, *Poa p* I, and *Phl p* I) in a therapeutic treatment regimen, are capable of modifying the allergic response of the individual to the allergen. Particularly, such preferred *Lol p* I peptides of the invention comprising at least one T cell epitope of *Lol p* I or at least two regions derived from *Lol p* I, each comprising at least one T cell epitope, when administered to an individual sensitive to ryegrass pollen are capable of modifying T cell response of the individual to the allergen, and they will thus be useful as therapeutics in addressing sensitivity to grasses.

A preferred isolated *Lol p* I peptide of the invention or portion thereof comprises at least one T cell epitope of *Lol p* I and accordingly, the peptide comprises at least approximately seven amino acid residues. For purposes of therapeutic effectiveness, preferred therapeutic compositions of the invention preferably comprise at least two T cell epitopes of *Lol p* I, and accordingly, the peptide comprises at least approximately eight amino acid residues and preferably at least fifteen amino acid

residues. Additionally, therapeutic compositions comprising preferred isolated peptides of the invention most preferably comprise a sufficient percentage of the T cell epitopes of the entire protein allergen so that a therapeutic regimen of administration of the composition to an individual sensitive to ryegrass pollen results in T cells of the individual being tolerized to the protein allergen. Synthetically produced peptides of the invention comprising up to approximately forty-five amino acid residues in length, and most preferably up to approximately thirty amino acid residues in length are particularly desirable, as increases in length may result in difficulty in peptide synthesis. Peptides of the invention may also be produced recombinantly as described above, and peptides exceeding 45 amino acids will be more easily produced recombinantly.

Peptides derived from the *Lol p* I protein allergen which exhibit T cell stimulatory properties and thus are believed to be useful therapeutics and/or intermediatea in developing tolerizing peptides comprise all or a portion of the following peptides: LPI-1 (SEQ ID NO: 3), LPI-1.1 (SEQ ID NO: 4), LPI-2 (SEQ ID NO: 5), LPI-3 (SEQ ID NO: 6), LPI-4 (SEQ ID NO: 7), LPI-4.1 (SEQ ID NO: 8), LPI-5 (SEQ ID NO: 9), LPI-6 (SEQ ID NO: 10), LPI-7 (SEQ ID NO: 11), LPI-8 (SEQ ID NO: 12), LPI-9 (SEQ ID NO: 13), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-12 (SEQ ID NO: 17), LPI-13 (SEQ ID NO: 19), LPI-14 (SEQ ID NO: 20), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-17 (SEQ ID NO: 24), LPI-18 (SEQ ID NO: 25), LPI-19 (SEQ ID NO: 26), LPI-20 (SEQ ID NO: 27), LPI-21 (SEQ ID NO: 28), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30) (Fig. 2) wherein the portion of the peptide preferably has a mean T cell stimulation index equivalent to, or greater than the mean T cell stimulation index of the corresponding peptide from which it is derived, as shown in Fig. 3. Even more preferably peptides derived from the *Lol p* I protein allergen comprise all or a portion of the following peptides: LPI-1.1 (SEQ ID NO: 4), LPI-2 (SEQ ID NO: 5), LPI-3 (SEQ ID NO: 6), LPI-4 (SEQ ID NO: 7), LPI-4.1 (SEQ ID NO: 8), LPI-8 (SEQ ID NO: 12), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-13 (SEQ ID NO: 19), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-19 (SEQ ID NO: 26), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29)and LPI-23 (SEQ ID NO: 30), as shown in Fig. 2. Additionally, even more preferred peptides derived from the *Lol p* I protein comprise the following peptides: LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID

NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30), all as shown in Fig. 2. Additional preferred peptides believed to T cell stimulating activity comprise the following peptides: LPI-16.2 (SEQ ID NO: 31), LPI-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID NO: 33), LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID NO: 35), LPI-16.7 (SEQ ID NO: 36), LPI-16.9 (SEQ ID NO: 37), LPI-16.10 (SEQ ID NO: 38), LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-18.8 (SEQ ID NO: 42), LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), LPI-20.6 (SEQ ID NO: 47), LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49), and LPI-23.4 (SEQ ID NO: 50).

One embodiment of the present invention features a peptide or portion thereof of *Lol p* I which comprises at least one T cell epitope of the protein allergen and has a formula X_n-Y-Z_m. According to the formula, Y is an amino acid sequence selected from the group consisting of LPI-1 (SEQ ID NO: 3), LPI-1.1 (SEQ ID NO: 4), LPI-2 (SEQ ID NO: 5), LPI-3 (SEQ ID NO: 6), LPI-4 (SEQ ID NO: 7), LPI-4.1 (SEQ ID NO: 8), LPI-5 (SEQ ID NO: 9), LPI-6 (SEQ ID NO: 10), LPI-7 (SEQ ID NO: 11), LPI-8 (SEQ ID NO: 12), LPI-9 (SEQ ID NO: 13), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-12 (SEQ ID NO: 17), LPI-13 (SEQ ID NO: 19), LPI-14 (SEQ ID NO: 20), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-17 (SEQ ID NO: 24), LPI-18 (SEQ ID NO: 25), LPI-19 (SEQ ID NO: 26), LPI-20 (SEQ ID NO: 27), LPI-21 (SEQ ID NO: 28), LPI-22 (SEQ ID NO: 29), LPI-23 (SEQ ID NO: 30), LPI-16.2 (SEQ ID NO: 31), LPI-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID NO: 33), LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID NO: 35), LPI-16.7 (SEQ ID NO: 36), LPI-16.9 (SEQ ID NO: 37), LPI-16.10 (SEQ ID NO: 38), LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-18.8 (SEQ ID NO: 42), LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), LPI-20.6 (SEQ ID NO: 47), LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49), and LPI-23.4 (SEQ ID NO: 50) and preferably selected from the group consisting of LPI-1.1 (SEQ ID NO: 4), LPI-2 (SEQ ID NO: 5), LPI-3 (SEQ ID NO: 6), LPI-4 (SEQ ID NO: 7), LPI-4.1 (SEQ ID NO: 8), LPI-8 (SEQ ID NO: 12), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-13 (SEQ ID NO: 19), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-19 (SEQ ID NO: 26), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), LPI-23 (SEQ ID NO: 30), LPI-16.2 (SEQ ID NO: 31), LPI-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID

NO: 33), LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID NO: 35), LPI-16.7 (SEQ ID NO: 36), LPI-16.9 (SEQ ID NO: 37), LPI-16.10 (SEQ ID NO: 38), LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-18.8 (SEQ ID NO: 42), LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), LPI-20.6 (SEQ ID NO: 47), LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49), and LPI-23.4 (SEQ ID NO: 50) and more preferably selected from the group consisting of LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), LPI-23 (SEQ ID NO: 30), LPI-16.2 (SEQ ID NO: 31), LPI-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID NO: 33), LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID NO: 35), LPI-16.7 (SEQ ID NO: 36), LPI-16.9 (SEQ ID NO: 37), LPI-16.10 (SEQ ID NO: 38), LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-18.8 (SEQ ID NO: 42), LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), LPI-20.6 (SEQ ID NO: 47), LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49), and LPI-23.4 (SEQ ID NO: 50), and most preferably selected from the group consisting of LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.2 (SEQ ID NO: 31), LPI-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID NO: 33), LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID NO: 35), LPI-16.7 (SEQ ID NO: 36), LPI-16.9 (SEQ ID NO: 37), LPI-16.10 (SEQ ID NO: 38), LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-18.8 (SEQ ID NO: 42), LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), LPI-20.6 (SEQ ID NO: 47), LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49), and LPI-23.4 (SEQ ID NO: 50). In addition, X_n are amino acid residues contiguous to the amino terminus of Y in the amino acid sequence of the protein allergen and Z_m are amino acid residues contiguous to the carboxy terminus of Y in the amino acid sequence of the protein allergen. In the formula, n is 0-30 and m is 0-30. Preferably, the peptide or portion thereof has a mean T cell stimulation index equivalent to greater than the mean T cell stimulation index of Y as shown in Fig. 3. Preferably, amino acids comprising the amino terminus of X and the carboxy terminus of Z are selected from charged amino acids, i.e., arginine (R), lysine (K), histidine (H), glutamic acid (E) or aspartic acid (D); amino acids with reactive side chains, e.g., cysteine (C), asparagine (N) or glutamine (Q); or amino acids with sterically small side chains, e.g.,

alanine (A) or glycine (G). Preferably n and m are 0-5; most preferably n + m is less than 10.

Another embodiment of the present invention provides peptides comprising at least two regions, each region comprising at least one T cell epitope of *Lol p* I and accordingly each region comprises at least approximately seven amino acid residues. These peptides comprising at least two regions can comprise up to 100 or more amino acid residues but preferably comprise at least about 14, even more preferably at least about 20, and most preferably at least about 30 amino acid residues of the *Lol p* I allergen. If desired, the amino acid sequences of the regions can be produced and joined by a linker to increase sensitivity to processing by antigen-presenting cells. Such linker can be any non-epitope amino acid sequence or other appropriate linking or joining agent. To obtain preferred peptides comprising at least two regions, each comprising at least one T cell epitope, the regions are arranged in the same or a different configuration from a naturally-occurring configuration of the regions in the allergen. For example, the regions containing T cell epitope(s) can be arranged in a noncontiguous configuration and can preferably be derived from the same protein allergen. Noncontiguous is defined as an arrangement of regions containing T cell epitope(s) which is different than that of the native amino acid sequence of the protein allergen from which the regions are derived. Furthermore, the noncontiguous regions containing T cell epitopes can be arranged in a nonsequential order (e.g., in an order different from the order of the amino acids of the native protein allergen from which the region containing T cell epitope(s) are derived in which amino acids are arranged from an amino terminus to a carboxy terminus). A peptide of the invention can comprise at least 15%, at least 30%, at least 50% or up to 100% of the T cell epitopes of *Lol p* I.

The individual peptide regions can be produced and tested to determine which regions bind immunoglobulin E specific for *Lol p* I and which of such regions would cause the release of mediators (e.g., histamine) from mast cells or basophils. Those peptide regions found to bind immunoglobulin E and to cause the release of mediators from mast cells or basophils in greater than approximately 10-15% of the allergic sera tested are preferably not included in the peptide regions arranged to form preferred peptides of the invention.

Examples of preferred peptide regions which do not bind to IgE (data not shown) include: LPI-1 (SEQ ID NO: 3), LPI-1.1 (SEQ ID NO: 4), LPI-2 (SEQ ID NO: 5), LPI-3 (SEQ ID NO: 6), LPI-4 (SEQ ID NO: 7), LPI-4.1 (SEQ ID NO: 8),

LPI-5 (SEQ ID NO: 9), LPI-6 (SEQ ID NO: 10), LPI-7 (SEQ ID NO: 11), LPI-8 (SEQ ID NO: 12), LPI-9 (SEQ ID NO: 13), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-12 (SEQ ID NO: 17), LPI-13 (SEQ ID NO: 19), LPI-14 (SEQ ID NO: 20), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-16.1 (SEQ ID NO: 23), LPI-17 (SEQ ID NO: 24), LPI-18 (SEQ ID NO: 25), LPI-19 (SEQ ID NO: 26), LPI-20 (SEQ ID NO: 27), LPI-21 (SEQ ID NO: 28), LPI-22 (SEQ ID NO: 29), LPI-23 (SEQ ID NO: 30), LPI-16.2 (SEQ ID NO: 31), LPI-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID NO: 33), LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID NO: 35), LPI-16.7 (SEQ ID NO: 36), LPI-16.9 (SEQ ID NO: 37), LPI-16.10 (SEQ ID NO: 38), LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-18.8 (SEQ ID NO: 42), LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), LPI-20.6 (SEQ ID NO: 47), LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49), and LPI-23.4 (SEQ ID NO: 50), the amino acid sequences of such regions being shown in Figs. 2 or 4, or portions of said regions comprising at least one T cell epitope.

Preferred peptides comprise various combinations of two or more of the above-discussed preferred regions, or a portion thereof. Preferred peptides comprising a combination of two or more regions (each region having an amino acid sequence as shown in Fig. 2 or Fig. 4), include the following:

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), and LPI-11 (SEQ ID NO: 15);

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), PLI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), and LPI-16 (SEQ ID NO: 22);

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), and LPI-16.1 (SEQ ID NO: 23);

LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), and LPI-16.1 (SEQ ID NO: 23);

LPI-10 (SEQ ID NO:14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), and LPI-20 (SEQ ID NO: 27);

LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29) and LPI-23 (SEQ ID NO: 30);

LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), and LPI-20 (SEQ ID NO: 27);

LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);

LPI-18 (SEQ ID NO: 25) and LPI-20 (SEQ ID NO: 27);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27) and LPI-23 (SEQ ID NO: 30);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27) and LPI-16.1 (SEQ ID NO: 23);

~~LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30) and LPI-16.1 (SEQ ID NO: 23);~~

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23) and LPI-11 (SEQ ID NO: 15);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23) and LPI-4.1 (SEQ ID NO: 8);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23), LPI-4.1 (SEQ ID NO: 8) and LPI-22 (SEQ ID NO: 29);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23), LPI-11 (SEQ ID NO: 15) and LPI-4.1 (SEQ ID NO: 8);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23), LPI-11 (SEQ ID NO: 15), LPI-4.1 (SEQ ID NO: 8) and LPI-22 (SEQ ID NO: 29);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-16.1 (SEQ ID NO: 23), LPI-22 (SEQ ID NO: 29) and LPI-23 (SEQ ID NO: 30); and
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-16.1 (SEQ ID NO: 23) and LPI-22 (SEQ ID NO: 29).

Additional preferred peptides comprising various combinations of two or more of the above discussed preferred regions include:
LPI-16.2 (SEQ ID NO: 31), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.3 (SEQ ID NO: 32), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.4 (SEQ ID NO: 33), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.5 (SEQ ID NO: 34), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.6 (SEQ ID NO: 35), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.7 (SEQ ID NO: 36), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.9 (SEQ ID NO: 37), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30); and
LPI-16.10 (SEQ ID NO: 38), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30).

In yet another aspect of the present invention, a composition is provided comprising at least two peptides (e.g., a physical mixture of at least two peptides), each comprising at least one T cell epitope of *Lol p* I. Such compositions can be in the form of a composition additionally with a pharmaceutically acceptable carrier of diluent for therapeutic uses, or with conventional non-pharmaceutical excipients for reagent use. When used therapeutically, an effective amount of one or more of such compositions can be administered simultaneously or sequentially to an individual sensitive to ryegrass pollen.

In another aspect of the invention, combinations of *Lol p* I peptides are provided which can be administered simultaneously or sequentially. Such combinations may comprise therapeutic compositions comprising only one peptide, or

more peptides if desired. Such compositions may be used simultaneously or sequentially in preferred combinations.

Preferred compositions and preferred combinations of *Lol p* I peptides which can be administered or otherwise used simultaneously or sequentially (comprising peptides having amino acid sequences shown in Fig. 2) include the following combinations:

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16 (SEQ ID NO: 22), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), and LPI-11 (SEQ ID NO: 15);

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), PLI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), and LPI-16 (SEQ ID NO: 22);

LPI-3 (SEQ ID NO: 6), LPI-4.1 (SEQ ID NO: 8), LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), and LPI-16.1 (SEQ ID NO: 23);

~~LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), and LPI-16.1 (SEQ ID NO: 23);~~

LPI-10 (SEQ ID NO:14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), and LPI-20 (SEQ ID NO: 27);

LPI-10 (SEQ ID NO: 14), LPI-11 (SEQ ID NO: 15), LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29) and LPI-23 (SEQ ID NO: 30);

LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), and LPI-20 (SEQ ID NO: 27);

LPI-15 (SEQ ID NO: 21), LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);

LPI-18 (SEQ ID NO: 25) and LPI-20 (SEQ ID NO: 27);

LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27) and LPI-23 (SEQ ID NO: 30);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27) and LPI-16.1 (SEQ ID NO: 23);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30) and LPI-16.1 (SEQ ID NO: 23);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23) and LPI-11 (SEQ ID NO: 15);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23) and LPI-4.1 (SEQ ID NO: 8);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23), LPI-4.1 (SEQ ID NO: 8) and LPI-22 (SEQ ID NO: 29);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23), LPI-11 (SEQ ID NO: 15) and LPI-4.1 (SEQ ID NO: 8);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-23 (SEQ ID NO: 30), LPI-16.1 (SEQ ID NO: 23), LPI-11 (SEQ ID NO: 15), LPI-4.1 (SEQ ID NO: 8) and LPI-22 (SEQ ID NO: 29);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-22 (SEQ ID NO: 29), and LPI-23 (SEQ ID NO: 30);
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-16.1 (SEQ ID NO: 23), LPI-22 (SEQ ID NO: 29) and LPI-23 (SEQ ID NO: 30); and
LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), LPI-16.1 (SEQ ID NO: 23) and LPI-22 (SEQ ID NO: 29).

Additional preferred compositions and preferred combinations of *Lol p* I peptides which can be administered or used simultaneously or sequentially (comprising peptides having amino acid sequences shown in Figs. 2 or 4) include the following combinations:

LPI-16.2 (SEQ ID NO: 31), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.3 (SEQ ID NO: 32), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);

LPI-16.4 (SEQ ID NO: 33), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.5 (SEQ ID NO: 34), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.6 (SEQ ID NO: 35), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.7 (SEQ ID NO: 36), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30);
LPI-16.9 (SEQ ID NO: 37), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30); and
LPI-16.10 (SEQ ID NO: 38), LPI-18 (SEQ ID NO: 25), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30).

In each of the above preferred compositions, peptides LPI-16.1 (SEQ ID NO: 23), LPI-18 (SEQ ID NO: 23), LPI-20 (SEQ ID NO: 27), and LPI-23 (SEQ ID NO: 30) may be substituted as follows: peptide LPI-16.1 (SEQ ID NO: 23) (Fig. 2) may be substituted with LPI-16.2 (SEQ ID NO: 31), LPI-16.3 (SEQ ID NO: 32), LPI-16.4 (SEQ ID NO: 33), LPI-16.5 (SEQ ID NO: 34), LPI-16.6 (SEQ ID NO: 35), LPI-16.7 (SEQ ID NO: 36), LPI-16.9 (SEQ ID NO: 37), and LPI-16.10 (SEQ ID NO: 38), all as shown in Fig. 4; peptide LPI-18 (SEQ ID NO: 25) (Fig. 2) may be substituted with peptides LPI-18.5 (SEQ ID NO: 39), LPI-18.6 (SEQ ID NO: 40), LPI-18.7 (SEQ ID NO: 41), LPI-18.8 (SEQ ID NO: 42) all as shown in Fig. 4; peptide LPI-20 (SEQ ID NO: 27) may be substituted with peptides LPI-20.2 (SEQ ID NO: 43), LPI-20.3 (SEQ ID NO: 44), LPI-20.4 (SEQ ID NO: 45), LPI-20.5 (SEQ ID NO: 46), and LPI-20.6 (SEQ ID NO: 47) all as shown in Fig. 4; peptide LPI-23 (SEQ ID NO: 30) may be substituted with peptides LPI-23.1 (SEQ ID NO: 48), LPI-23.2 (SEQ ID NO: 49) and LPI-23.4 (SEQ ID NO: 50), all as shown in Fig. 4.

The present invention is further illustrated by the following non-limiting Figures and Examples.

EXAMPLES

Example 1 - Isolation and Cloning of Nucleic Acid Sequence Coding for *Lol p* I

Total mRNA was extracted from mature ryegrass pollen by the phenol method of Herrin and Michaels, *supra*. Double-stranded cDNA was synthesized from 1µg of

total mRNA using a commercially available kit (cDNA SYNTHESES SYSTEM PLUS KIT, BRL, Gaithersburg, MD). After a phenol extraction and ethanol precipitation, the cDNA was blunted with T4 DNA polymerase (Promega, Madison, WI) and ligated to ethanol-precipitated, self-annealed AT and AL oligonucleotides for use in a modified Anchored PCR reaction, according to the method in Rafnar *et al.* (1991), *J. Biol. Chem.*, **266**: 1229-1236; Frohman *et al.* (1990), *Proc. Natl. Acad. Sci. USA*, 85:8998-9002; and Roux *et al.* (1990), *BioTech.*, **8**: 48-57. Oligonucleotide AT has the sequence 5'-GGGTCTAGAGGTACCGTCCGATCGATCATT-3' (SEQ ID NO: 71) (Rafnar *et al. supra*). Oligonucleotide AL has the sequence AATGATCGATGCT (SEQ ID NO: 72) (Rafnar et al. *supra.*).

Polymerase chain reactions (PCR) were carried out using a commercially available kit (GeneAmp® DNA Amplification kit, Perkin Elmer Cetus, Norwalk, CT) whereby 10 µl 10x buffer containing dNTPs were mixed with 1 µg each of primer AP, which has the sequence 5'-GGGTCTAGAGGTACCGTCCG-3' (SEQ ID NO: 73) (Rafner et al. *supra.*) and LpA-5, which has the sequence 5'-CCCTGCAGATTATTTGAGATCTTGAG-3' (SEQ ID NO: 74), cDNA (3-5 µl of a 20 µl linkered cDNA reaction mix), 0.5 µl Amplitaq DNA polymerase, and distilled water to 100 µl.

Nucleotides 1 through 8 (5'-CCCTGCAG) of LpA-5 correspond to a Pst I site added for cloning purposes; the remaining nucleotides correspond to the non-coding strand sequence complementary to nucleotides 483 through 500 as shown in Fig. 6.

The samples were amplified with a programmable thermal controller (MJ Research, Inc., Cambridge, MA). The first 5 rounds of amplification consisted of denaturation at 94°C for 1 minute, annealing of primer to the template at 45°C for 1.5 minutes, and chain elongation at 70°C for 2 minutes. The final 20 rounds of amplification consisted of denaturation as above, annealing at 55°C for 1.5 minutes, and elongation as above. Five percent (5 µl) of this initial amplification was then used in a secondary amplification whereby 10 µl 10x buffer containing dNTPs was mixed with 1 µg each of primer AP and primer LpA-3, which has the sequence 5'-CCCTGCAGTCATGCTCACTTGGCCGAGTA-3' (SEQ ID NO: 75), 0.5 µl Amplitaq DNA polymerase, and distilled water to 100 µl. The secondary PCR reaction was performed as described herein. Nucleotides 1 through 8 (5'-CCCTGCAG-3') of LpA-3 correspond to a Pst I site added for cloning purposes; nucleotides 9 through 12 (5'-TCA-3') correspond to the complementary sequence for a new stop codon, and the remaining nucleotides correspond to the non-coding strand sequence complementary

to nucleotides 793 through 810 of the full length clone of *Lol p* I as shown in Fig. 1, including translated sequence of *Lol p* I (Fig. 1), the native stop codon and 3' untranslated sequence.

Amplified DNA was recovered by sequential chloroform, phenol, and chloroform extractions, followed by precipitation at -20°C with 0.5 volumes of 7.5 ammonium acetate and 1.5 volumes of isopropanol. After precipitation and washing with 70% ethanol, the DNA was simultaneously digested with *Xba* I and *Pst* I in a 15 µl reaction and electrophoresed through a preparative 3% GTG NuSieve low melt gel (FMC, Rockport, ME). The appropriate sized DNA band was visualized by EtBr staining, excised, and ligated into appropriately digested M13mp18 for sequencing by the dideoxy chain termination method (Sanger *et al.* (1977), *Proc. Natl Acad Sci USA*, 74: 5463-5476) using a commercially available sequencing kit (Sequenase kit, U.S. Biochemicals, Cleveland, OH).

Both strands were sequenced using M13 forward and reverse primers (N.E. BioLabs, Beverly, MA) and internal sequencing primers LpA-13, LpA-12, LpA-9, LpA-2, LpA-7, LpA-10, and LpA-IA. LpA-13 has the sequence 5'-GAGTACGGCGACAAGTGGC-3' (SEQ ID NO: 76), which corresponds to nucleotides 121 through 139 as shown in Fig. 1. LpA-12 has the sequence 5'-TTCGAGATCAAGTGCACC-3' (SEQ ID NO: 77), which corresponds to nucleotides 310 through 318 as shown in Fig. 1. LpA-9 has the sequence 5'-GTGACAGCCTCGCCGG-3' (SEQ ID NO: 78), which corresponds to the non-coding strand sequence complementary to nucleotides 335 through 350 as shown in Fig. 1. LpA-2 has the sequence 5'-GGGAATTCCATGGCGAAGAAGGGC-3' (SEQ ID NO: 79). Nucleotides 1 through 7 (5-GGGATT-3') of LpA 2 correspond to part of an Eco-RI restriction site added for cloning purposes; the remaining sequence of LpA-2 corresponds to nucleotides 425 through 441 of Fig. 1. LpA-7 has the sequence 5'-GTGCCGTCCGGGTACT-3' (SEQ ID NO: 80), and corresponds to non-coding strand sequence complementary to nucleotides 503 through 518 of Fig. 1. LpA-10 has the sequence 5'-CCGTCGACGTACTTCA-3' (SEQ ID NO: 81), which corresponds to non-coding strand sequence complementary to nucleotides 575 through 590 of Fig. 1. LpA-IA has the sequence 5'-GGAGTCGTGGGGAGCAGTC-3' (SEQ ID NO: 82), which corresponds to nucleotides 654 through 672 of Fig. 1.

Multiple clones from several independent PCR reactions were sequenced. The nucleotide (SEQ ID NO: 1) and deduced amino acid sequences (SEQ ID NO: 2) of a representative clone of *Lol p I*, clone 26.*j* are shown in Fig. 1. As shown in Fig.1, the

nucleic acid sequence coding for *Lol p I* has an open reading frame beginning with an ATG initiation codon at nucleotides 16-18 ending with a TGA stop codon at nucleotides 805-807. The translated protein has a deduced amino acid sequence of 263 amino acids with a predicted molecular weight of 28.4 kD and a pI of 5.55. The initiating methionine is numbered amino acid -23, with amino acid numbered +1 corresponding to the NH_2-terminus of the mature protein, as defined by amino acid sequencing (Cottam *et al.* (1986), *Biochem. J.*, **234**: 305-310). Amino acids -23 through -1 (Fig.1), correspond to a leader sequence that is cleaved to yield the mature protein; the mature protein is therefore composed of 240 amino acids and has a predicted molecular weight of 26.1 kD and a pI of 5.38. There is a single potential N-linked glycosylation site at amino acid 9.

Amino acids 1 through 30 of clone 26.*j* (Fig. 1) correspond exactly to the published sequence of the NH_2 terminus of *Lol p* I (Cottam *et al.*, *supra*). Amino acids 213 through 240 of clone 26.*j* (Fig. 1) correspond exactly to the published internal amino acid sequence of *Lol p* I (Esch and Klapper (1989), *Mol. Immunol.*, **26**: 557-561).

Example 2 - Identification of Polymorphisms in *Lol p I*

A number of polymorphisms in the nucleotide sequence coding for *Lol p* I were discovered during the amplification and sequencing of different *Lol p* I clones. Some of the polymorphisms cause an amino acid change relative to that of clone 26.*j*, while others are silent polymorphisms that do not cause an amino acid change. The polymorphisms found in the sequence coding for *Lol p* I are summarized in Table 1. The nucleotide base numbers are those of the sequence of clone 26.*j* shown in Fig 1.

Table 1
Polymorphisms Detected in *Lol p I*

	Nucleotide Polymorphism	Amino Acid Polymorphism
1	GGC_{215}→GGA/GGT	None
2	$G_{234}AC_{236}$→GAT	D_{45}→N
3	GTT_{239}→GTC	None
4	CGT_{351}→CGC	None
5	GGC_{356}→GGT	None
6	AAC_{389}→AAT	None
7	CCC_{396}→CCT	None
8	CAT_{413}→CAC	None
9	GCC_{434}→GCA	None
10	GAC_{530}→GAT	None
11	$GG_{532}C$→GAC	G_{144}→D
12	CCG_{542}→CCA	None
13	ACA_{545}→ACG	None
14	$GC_{562}T$→GGT	A_{154}→G
15	CTC_{581}→CTG	None
16	GCG_{626}→GCC	None
17	ATC_{782}→ATT	None
18	CCT_{785}→CCC	None

All confirmed nucleotide polymorphisms (polymorphisms observed in the sequence analysis of clones from two independent PCR reactions) are shown relative to the sequence of clone 26.*j* (Fig.1) (SEQ ID NO: 1). The polymorphic residues in their respective codon triplets are numbered. Productive amino acid changes are also shown; most nucleotide polymorphisms are silent and do not result in an amino acid change. Twenty-eight potential polymorphisms have only been observed in clones from single PCR reactions. Seventeen of these 28 potential polymorphisms are silent mutations and do not result in an amino acid polymorphism; the remaining 11 potential polymorphic sites would result in the following amino acid changes, specifically: T_{11}

$\rightarrow$M, $A_{49}\rightarrow$V, $R_{67}\rightarrow$S, $K_{79}\rightarrow$R, $V_{90}\rightarrow$I, $Q_{133}\rightarrow$R, $I_{162}\rightarrow$T, $V_{173}\rightarrow$E, $I_{187}\rightarrow$T, $V_{223}\rightarrow$F and $K_{232}\rightarrow$R. The potential polymorphism at amino acid 223 ($V_{223}\rightarrow$F) has been previously reported. (Perez *et al.*, *supra*)

Example 3 - Human IgE reactivity to Purified Recombiant and Native *Lol p* I

Cloned DNA encoding *Lol p* I and *Lol p* IX was expressed in E. coli and purified on a Ni-chelating affinity column. Monoclonal antibodies were also used to affinity purify and distinguish isoforms of these and native grass proteins. The recombinant *Lol p* I was compared to biochemically purified native *Lol p* I and *Lol p* IX in mAb and human IgE reactivity studies (data not shown). The reactivity of human IgE to the recombinant and native forms was equivalent when measured by direct binding ELISA. In competition assays, the native *Lol p* I and *Lol p* IX proteins could completely inhibit IgE binding to *Lol p* soluble pollen extract (SPE), whereas the recombinant form of *Lol p* I and *Lol p* IX could only partially inhibit IgE binding to the extract. However, the recombinant *Lol p* I and *Lol p* IX was still active in these competition assays. These asays were then extended to western blot inhibition studies; both methods confirm the previous finding that Group I and Group IX constitute one of the major allergenic proteins of *Lolium perenne* grass pollen. Furthermore, the *Lol p* I and *Lol p* IX native and recombinant allergens showed ihibition of grass allergic patient IgE binding to soluble pollen extracts of other grass species (*Dac g*, *Phl p* and *Poa p*). The degree to which *Lol p* I and *Lol p* IX proteins successfully compete for IgE binding to these other grasses implies a hierarchy of homology between the species. These studies confirm and extend the findings of shared IgE epitopes between temperate grass allergens.

The procedures used for the foregoing examples were as follows:

Extraction and Depigmentation of Allergens

Defatted *Lol p* I pollen was extracted twice, overnight at 4°C in 50mM phosphate buffer, 15mM NaCl, pH 7.2 and protease inhibitors (PMSF, Luepeptin, SPTI and pepstatin). The extract was then depigmented by batch absorption with DE-52 (Whatman) in 50mM phosphate buffer, 0.3M NaCl, pH 7.2.

Biochemical Purification of *Lol p* I Allergen

Depigmented *Lol p* I extract was dialyzed into H_2O, pH 8.0 by addition of NH_4OH. This mateial was loaded onto a DE-52 column and eluted stepwise with 1mM, 4.5mM and 7.5mM NaH_2PO_4. The majority of the Group I allergens was eluted with 4.5mM NaH_2P0_4. A further separation of Group I was accomplished by running this DE-52 enriched fraction over A (26/60) superdex 75 column (Pharmacia).

Immunoaffinity Purification of *Lol p* IX Allergen

1B9 ascites was precipitated by 50% $(NH_4)_2SO_4$, followed by purification over Q-sepharose (Pharmacia). Purified 1B9, an anti-*Lol p* IX antibody, was then coupled to Affigel-10 (Biorad), according to the manufacturer's instructions. Either depigmented pollen extract or DE-52 enriched material was circulated over the 1B9 affigen column overnight at 4°C. The column was washed with PBS, PBS + 0.5M MaCl and then eluted with 0.1M Glycine, pH2.7. Eluted *Lol p* IX fractions were neutralized with 1M tris-base, pH 11.

Expression and Purification of Recombinant *Lol p* I

Lol p I cDNA's encoding from the first amino acid of the mature protein to the stop codon were ligated into pET11dΔHR containing a leader which encoded 6 histidines. The HIS_6 was used for purification over a nickel-NTA agarose column (Qiagen). r*Lol p* I was expressed in *E. coli*.

SDS-PAGE, Electroblotting and Immunoblotting

Electrophoresis was performed using 12.5% polyacrylamide gels. The samples were run under reducing conditions (4 hours at 40mA constant current). After electrophoresis the protein was transferred to nitrocellulose membrane (1.5 hours at 1.5A). The blots were stained with 1% India ink, and then blocked with 1% defatted milk, 1% FCS in Tween solution (2mM Tris-HCl pH 7.5, 0.71M NaCl, and 0.05% Tween 20) for 1 hour. The human plasma samples were pre-absorbed with blank

nitrocellulose for 1.5 hours prior to incubation. Blot sections were incubated with 1st antibodies diluted in 1% milk/Tween solution overnight at room temperature (RT). The blot sections were washed three times and inucbated in the appropriate biotinylated 2nd AB (1:2500) for 2 hours at RT. The blot sections were washed three times and finally incubated with ^{125}I-streptavidin 1 hour at RT. The sections were washed extensively to remove unbound label and exposed to film. Autoradiography was carried out at -80°C.

Direct, Competition and Depletion ELISA

Microtiter plates were coated with 2.5-10.0µg/mL of coating antigen (grall soluble pollen extract (SPE), *Lol p* I, *Lol p* IX, *Lol p* IX, recombinant *Lol p* I, and/or recombinant *Lol p* IX) in PBS at 100µL/well and incubated overnight at 4°C. The plates were washed three times between each step with PBS-T (Phosphate buffered saline + 0.05% Tween 20). The unbound antigen was removed and the plate blocked with 300µL/well of 1MG/ML PVP in 0.5% gelatin/PBS for one hour at room temperature (RT). All subsequent reagents were added at 100µL/well for direct ELISA, serially diluted human plasma was added to duplicate wells and incubated overnight at 4°C. This was followed by biotinylated goat anti-human IgE (1:1,000) for 1 hour at RT, then streptavidin-HRPO (1:10,000) for 1 hour at RT. TMB substrate and H_2O_2 were freshly mixed and added; the color was allowed to develop for 2-5 minutes. The reaction was stopped by the addition of 1M phosphoric acid. The plates were read on a dynatech plate reader at 450NM and the absorbances of duplicate wells were averaged.

For the competition ELISA, the human plasma samples were mixed with an equal volume of serially diluted antigen or with PBS-T (as a control). These samples were incubated overnight at 4°C before addition to the microtiter plate and performing the remaining steps of the ELISA as stated above.

For the depletion ELISA, the human plasma was pre-incubated on antigen or PBS coated wells, collected and re-incubated on freshly coated wells. The ELISA was then performed as outlined above.

EXAMPLE 4 - Human T Cell Studies with *Lol p* I

Synthesis of Overlapping Peptides

Ryegrass *Lol p* I overlapping peptides were synthesized using standard Fmoc/tBoc synthetic chemistry and purified by Reverse Phase HPLC. Fig. 2 shows *Lol p* I peptides used in these studies (SEQ ID NO: 3-30). The peptide names are consistent throughout.

IgE Binding Studies with overlapping peptides

None of the peptides shown in Fig. 2 bound a detectable amount of IgE from pooled human plasma when analyzed in a solid phase ELISA assay (data not shown). The procedure for the ELISA assay with the overlapping peptides was substantially the same as that described in Example 3.

T Cell Responses to Ryegrass Antigen Peptides

Peripheral blood mononuclear cells (PBMC) were purified by lymphocyte separation medium (LSM) centrifugation of 60 ml of heparinized blood from grass-allergic patients who exhibited clinical symptoms of seasonal rhinitis and were MAST and/or skin test positive for grass. Long-term T cell lines were established by stimulation of $2x10^6$ PBL/ml in bulk cultures of complete medium IRPMI-1640, 2 mM L-glutamine, 100 U/ml penicillin/streptomycin, $5x10^{-5}$M 2-mercaptoethanol, and 10 mM HEPES, supplemented with 5% heat-inactivated human AB serum) with 25 mg/ml of purified native *Lol p* I (95% pure with a single band on protein gel) for 6 days at 37°C in a humidified 5% CO_2 incubator to select for *Lol p* I reactive T Cells. This amount of priming antigen was determined to be optimal for the activation of T cells from most grass-allergic patients. Viable cells were purified by LSM centrifugation and cultured in complete medium, supplemented with 5 units recombinant human IL-2/ml and 5 units recombinant human IL-4/ml for up to 3 weeks until the cells no longer responded to lymphokines and were considered "rested." The ability of the T cells to proliferate to selected peptides, recombinant *Lol p* I (r*Lol p I*), purified native *Lol p* I, recombinant *Lol p* IX (r*Lol p* IX), or *Der p* I (r*Der p* I) was then assessed. For assay, $2x10^4$ rested cells were restimulated in the presence of $2x10^4$ autologous Epstein-Barr virus (EBV)-transformed B cells (prepared as described below) with 2-50 mg/ml of r*Lol p* I, purified native *Lol p* I, r*Der p* I, or r*Lol p* IX, in a volume of 200 ml complete medium in duplicate wells in 96-well round-

bottom plates for three days. Each well then received 1 mCi tritiated thymidine for 16-20 hours. The counts incorporated were collected onto glass fiber filter mats and processed for liquid scintillation counting. The varying antigen dose in assays with r*Lol p* I, purified native *Lol p* I, and recombinant *Lol p* IX and several antigenic peptides (i.e., peptides that induce an immune response, or, specifically, a positive T cell response in these assays) synthesized as described above were determined. The titrations were used to optimize the dose of peptides in T cell assays. The maximum response in a titration of each peptide is expressed as the stimulation index (S.I.). The S.I. is the counts per minute (CPM) incorporated by cells in response to peptide, divided by the CPM incorporated by cells in medium only. An S.I. value equal to or greater than 2 times the background level is considered "positive" and indicates that the peptide contains a T cell epitope. The positive results were used in calculating mean stimulation indices for each peptide for the group of patients tested. The results (not shown) demonstrate that one patient responds well to r*Lol p* I and purified native *Lol p* I, as well as to *Lol p* I peptides but not to recombinant *Der p* I. This indicated that *Lol p* I T cell epitopes are recognized by T cells from this particular allergic patient and that r*Lol p* I contains such T cell epitopes. T cells from the majority of patients also reacted to r*Lol p* IX, suggesting a presence of *Lol p* IX antigen in the purified native *Lol p* I prep that was used to prime T cells.

The above procedure was followed with a number of other patients. Individual patient results were used in calculating the mean S.I. for each peptide if the patient responded to the *Lol p* I protein at an S.I. of 2.0 or greater and the patient responded to at least one peptide derived from *Lol p* I at an S.I. of 2.0 or greater. A summary of positive experiments from 35 patients is shown in Fig. 3. All 35 T cell lines responded to purified native *Lol p* I and r*Lol p* I. The numbers enclosed in the parentheses denote percentage of patients responding to that particular peptide. The bar represents the positivity index for each peptide (% of patients responding multiplied by mean S.I.).

Preparation of EBV-transformed B Cells for Use as Antigen-presenting Cells

Autologous EBV-transformed cell lines were derived by incubating $5x10^6$ PBL with 1 ml of B-59/8 Marmoset cell line (ATCC CRL1612, American Type Culture Collection, Rockville, MD) conditioned medium in the presence of 1 mg/ml phorbol 12-myristate 13-acetate (PMA) at 37°C for 60 minutes in 12x75 mm polypropylene round-bottom Falcon snap cap tubes (Becton Dickinson Labware, Lincoln Park, NJ).

These cells were then diluted to 1.25×10^6 cells/ml in the RPMI-1640 medium that was supplemented with 10% head-inactivated fetal bovine serum in place of the 5% human AB serum and cultured in 200 ml aliquots in flat-bottom culture plates until visible colonies were detected. They were then transfered to larger wells until the cell lines were established.

Those skilled in the art will appreciate that the invention described is susceptible to variations and modification other than those specifically described. It is understood that the invention includes all such variations and modifications. The invention also includes all steps, features, compositions and compounds referred to or indicated in this specification, individually or collectively, and any and all combinations of any two or more of said steps or features.

Example 5 - Cloning and Expression of *Dac g* I, *Poa p* I and *Phl p* I

A. Cloning of *Dac g* 1.

RNA was obtained from the pollen of *Dactylis glomerata* using a standard acid phenol extraction procedure (Sambrook *et al.* (1989), *Molecular Cloning: A laboratory manual. 2nd Edition.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). This and other pollens described below were purchased from Greer Laboratories (Lenoir, NC). Single and double stranded cDNA was prepared from total *D. glomerata* RNA using the BRL cDNA Synthesis System (Gaithersberg, MD), blunted using standard procedures (Sambrook et al. (1989) *supra*), and ligated to self-annealed oligonucleotides AT (5'-GGGTCTAGAGGTACCGTCCGATCGATCATT-3') (SEQ ID NO: 71) and AL (5'-AATGATCGATGCT-3') (SEQ ID NO: 72) (Rafnar *et al.* (1991), *J. Biol. Chem.*, **266**:1229-1236).

The amino portion of the gene encoding *Dac g* 1, including 5' untranslated sequence, nucleotide sequence encoding the predicted leader sequence and nucleotide sequence encoding the first portion of the mature protein, was cloned using the polymerase chain reaction (PCR). Oligonucleotide primers AP-2 (5'-GGGTCTAGAGGTACCGTCC-3') (SEQ ID NO: 83) and LpA-7 (5'-GTGCCGTCCGGGTACT-3') (SEQ ID NO: 80) were used in a primary amplification. Oligonucleotide primers AP-2 and LpA-9 (5'-GTGACAGCCTCGCCGG-3') (SEQ ID NO: 78) were used in a secondary amplification using 10% of the primary amplification as template cDNA. PCRs were carried out using the GeneAmp DNA Amplification kit (Perkin Elmer, Norwalk, CT)

using a programmable thermal controller from MJ Research, Inc. (Cambridge, MA). Samples were amplified for 24 cycles by heating to 94°C for 1 min, 54°C for 1.5 min and 70°C for 1 min.

The resulting PCR product was blunted with T4 DNA polymerase (Sambrook *et al.* (1989) *supra*) and digested with the restriction endonuclease *Xba*I. Unless otherwise stated, all endonucleases and polymerases were obtained from New England BioLabs (Beverly, MA). A band of approximately 400 base pairs was isolated from a low melting temperature agarose gel (FMC, Rockland, ME) and ligated into appropriately digested pUC19. The clones 22.2 and 22.5 were subsequently identified by dideoxysequencing (Sanger et al. (1977), *Proc. Natl. Acad. Sci. USA*, 74:5460-5463) to contain nucleotide sequence of the gene encoding *Dac g* 1.

A 600 base pair cDNA containing internal nucleotide sequence of the gene encoding *Dac g* 1 was amplified using the primers DGI-3 (5'-TTGGATCCTACGGCAAGCCGACCGGC-3') (SEQ ID NO: 84) and LpA-10 (5-CCGTCGACGTACTTCA-3') (SEQ ID NO: 81). A 300 base pair cDNA containing internal *Dac g* 1 sequence was amplified using the primers DGI-4 (5'-TTGGATCCATCCCGAAGGTGCCCCCGGG-3' (SEQ ID NO: 85), wherein G at position 14 can also be A) and LpA-9 (5'-GTGACAGCCTCGCCGG-3') (SEQ ID NO: 78). The cDNAs were amplified for 34 cycles by heating to 94°C for 45 sec, 60°C for 45 sec and 72°C for 1 min. These PCR products were blunted with T4 DNA polymerase as above, digested with *BamH*I and ligated into appropriately digested pUC19. Clones 86.1 (600 base pairs) and 88.6 (300 base pairs) were sequenced and found to contain sequence of the gene encoding *Dac g* 1.

The carboxy portion of the gene encoding *Dac g* 1, including the 3' untranslated region, was cloned using oligonucleotide primers AP (5'-GGGTCTAGAGGTACCGTCCG-3') (SEQ ID NO: 73) and DGI-8 (5'-AGGTGACCTTCCACGTCG-3') (SEQ ID NO: 86) in a primary PCR and oligonucleotide primers AP and DGI-9 (5'-TTGGATCCTGGCGCTGCTGGTGAAGTA-3') (SEQ ID NO: 87) in a secondary PCR. Material was amplified for 25 cycles of heating to 94°C for 1 min, 60°C for 40 sec and 74°C for 1 min. The 700 base pair PCR product was digested with *BamH*I and *Asp*718 (Boehringer Mannheim, Indianapolis, IN), isolated and digested into appropriately digested pUC19 as described above. The clones 119.2, 119.4, 119.6, 119.9 and 119.12 were isolated, sequenced and found to contain sequence of the gene encoding *Dac g* 1.

cDNA clones encoding the mature *Dac g* 1 protein were obtained by PCR with the oligonucleotide primers DGI-7Eco (5'-TTGAATTCATCCCGAAGGTGCCCCCG-3' (SEQ ID NO: 88), wherein G at position 14 can also be A) and PhA-1.2 (5'-TTGGTACCTCACTTGGACTCGTAGCT-3') (SEQ ID NO: 89). The cDNAs were amplified for 24 cycles of heating to 94°C for 1 min, 54°C for 1.5 min and 70°C for 1 min. The amplified cDNA was digested with *EcoR*I and *Asp*718, isolated, and ligated into the appropriately digested pUC19. The cDNA clones 106.5, 106.6, 106.9 and 106.12 were identified as containing *Dac g* 1 sequence by dideoxysequencing. The nucleotide (SEQ ID NO: 51) and deduced amino acid (SEQ ID NO: 52) sequences of clone 106.5 are shown in Fig. 5. Nucleotides 509-515 (encoding amino acids 171 and 172) are from the sequence of clone 106.12. The sequence of clone 106.5 was not resolved in this region.

The insert from clone 106.5 was isolated and ligated into appropriately digested expression vector pET-11d (Novagen, Madison, WI: Jameel et al. (1990), *J. Virol.*, **64**:3963-3966). The pET-11d vector had been modified to contain a sequence encoding 6 histidines (His 6) immediately 3' of the ATG initiation codon followed by a unique *EcoR*I endonuclease restriction site. A second *EcoR*I endonuclease restriction site in the vector, along with neighboring *Cla*I and *Hin*dIII endonuclease restriction sites, had previously been removed by digestion with *EcoR*I and *Hin*dIII, blunted and religated.

A recombinant clone was used to transform *Escherichea coli* strain BL21-DE3. A culture was grown to A600 of 1.0, IPTG was added to 1 mM final concentration and grown for an additional 2 hours. Bacteria was recovered by centrifugation (7,930 G, 10 min) and lysed in 90 ml of 6 M Guanidine-HCl, 0.1 M Na_2HPO_4, pH 8.0 for 1 hour with vigorous shaking. The recombinant *Dac g* 1 was purified from the extract on a Ni^{+2} chelating column (Hochuli et al. (1987) *J. Chromatog.*, **411**:177-184; Hochuli et al. (1988), *Bio/Tech.*, **6**:1321-1325).

B. Cloning of *Poa p* I.

RNA was isolated from the pollen of *Poa pratensis*, double stranded cDNA was prepared and self-annealed oligonucleotides AT and AL were added as described in section A, above. PCR product was amplified using oligonucleotide primers Phl-7 (5'-CCGAATTCGTGGAGAAGGGGTCCAA-3') (SEQ ID NO: 90) and Poa-1 (5'-TTAGGATCCTCACTTATCATAIGACGTATC-3' (SEQ ID NO: 91), wherein C at

position 13 can also be T, A at position 16 can also be G, A at position 19 can also be G, G at position 23 can also be C, A at position 24 can also be T, C at position 25 can also be T or A or G and A at position 28 can be G). All *Poa p* 1 clones were amplified by 20 cycles of heating to 94°C for 1 min, 55°C for 1 min and 72°C for 1 min. The amplified material was finally heated to 72°C for 5 min. Three clones, 11, 15 and 17, were isolated that contained part of the nucleotide sequence for the gene that encodes *Poa p* 1. The *Dac g* 1 sequence encoded by clones 11, 15 and 17 corresponds to amino acids 151 - 240 of Fig. 6.

Clones containing partial nucleotide sequences of the gene encoding *Poa p* 1 were derived from PCRs that used oligonucleotide primers AP and Poa-3 (5'-TTGAATTCCTTGTCATTGCCCTTCTG-3') (SEQ ID NO: 92) in the primary PCR and AP and Poa-4 (5'-AAGAATTCCTTCTGCTTGATGTCCAC-3') (SEQ ID NO: 93) in the secondary PCR. Other clones were derived from PCRs that used oligonucleotide primers AP and Poa-6 (5'-ATGAATTCGAGTCGTGGGGAGCCGTC-3') (SEQ ID NO: 94) in the primary PCR and AP and Poa-7 (5'-ATGAATTCGTCTGGAGGATCGACACC-3') (SEQ ID NO: 95) in the secondary PCR. Clones 58, 59 and 63 were derived from the PCR using primers AP and Poa-4. Clones 91 and 97 were derived from the PCR using primers AP and Poa-7.

Additional clones were derived from a PCR that used oligonucleotide primers Poa-1 and Poa-5 (5'-ATGAATTCATCGCAAAGGTTCCCCCC-3' (SEQ ID NO: 96), wherein A at position 14 can also be G or C or T). These clones, 113, 114 and 115, corresponded to the portion of the gene that encoded amino acids 1 - 240 of *Poa p* 1 (see Fig. 6). The nucleotide (SEQ ID NO: 53) and deduced amino acid (SEQ ID NO: 54) sequences of clone 114 are shown in Fig. 6. Nucleotide 93 in Fig. 6 was not resolved and could be a G or a C or a T or an A and is represented by the letter "N". Nucleotide 94 in Fig. 6 was not conclusively resolved and could be a G or a C or a T but not an A and is represented by the letter "B". The codon containing nucleotide 93 (GGN) encodes a Glycine at residue 31. The codon containing nucleotide 94 (BCC) encodes an Alanine (GCC), a Proline (CCC), or a Serine (TCC) at amino acid 32. The amino acid at residue 32 in Fig. 6 is represented by an "X".

Inserts from clones 11 and 114 were isolated and ligated into appropriately digested expression vector pET-11d (Novagen, Madison, WI: Jameel et al. (1990) *J. Virol.* 64:3963-3966). Recombinant proteins were expressed as descibed in section A, above.

C. Cloning of *Phl p* 1.

RNA was isolated from the pollen of *Phleum pratense*, double stranded cDNA was prepared and self-annealed oligonucleotides AT and AL were added as described in section A, above. Clones were derived from a PCR that used oligonucleotide primers PhA1.1 (5'-TTTGGATCCTCACTTGGACTCGTAGCT-3') (SEQ ID NO: 97) and Phl-2 (5'-TTGAATTCTCGCGAAGGTGCCCCCG-3' (SEQ ID NO: 98), wherein G at position 13 can also be A). These clones, 20 and 22, corresponded to the portion of the gene that encoded amino acids 1 - 240 of *Phl p* 1 (see Fig. 7). The nucleotide (SEQ ID NO: 55) and deduced amino acid (SEQ ID NO: 56) sequences of clone 20 are shown in Fig. 7.

Clones containing partial nucleotide sequence of the gene encoding *Phl p* 1 were derived from a PCR using oligonucleotide primers Phl-7 (5'-CCGAATTCGTGGAGAAGGGGTCCAA-3') (SEQ ID NO: 90) and PhA1.1. Clones 47-52 were derived from this PCR. These clones encoded amino acids 151 through 240 of Fig. 7.

Inserts from clones 22 and 51 were isolated and ligated into appropriately digested expression vector pET-11d (Novagen, Madison, WI: Jameel et al. (1990) *J. Virol.* 64:3963-3966). Recombinant proteins were expressed as descibed in section A, above.

Example 6 - Comparison of *Dac g* 1, *Phl p* 1 and *Poa p* 1 With *Lol p* 1.

The sequences for *Dac g* 1 (Fig. 5) (SEQ ID NO: 58), *Phl p* 1 (Fig. 7) (SEQ ID NO: 59) and *Poa p* 1 (Fig. 6) (SEQ ID NO: 60) were compared with *Lol p* 1 (SEQ ID NO: 57). The amino acid sequences of these Group 1 allergens had 95% (*Dac g* 1), 91% (*Phl p* 1) and 91% (*Poa p* 1) identity, respectively, with *Lol p* 1. This comparison is shown schematically in Fig. 8. The complete sequence of *Lol p* 1 is shown in standard one letter code. Only differences from the *Lol p* 1 sequence are shown for the other Group 1 allergens; identity is indicated by a dash (-). Potential amino acid polymorphisms were predicted by detected nucleotide polymorphisms in each sequence. Such potential polymorphisms are shown by superscript and subscript letters at the site of the polymorphism.

T cell epitope containing peptides of *Lol p* 1, peptides 16.1 (SEQ ID NO: 23), 18 (SEQ ID NO: 25), 20 (SEQ ID NO: 27) and 23 (SEQ ID NO: 30), were defined in

Example 4 (Fig. 3). The sequences of the other Group 1 allergens are very conserved in these regions. Since the Group 1 allergens are homologous, the major T cell epitope containing peptides of *Lol p* 1 are likely to be the major T cell epitope containing regions in the related grasses. Comparison of the sequences of the *Lol p* 1 peptides with the homologous peptides containing *Dac g* 1, *Phl p* 1 and *Poa p* 1 polymorphisms are shown in Fig. 9 (SEQ ID NO: 23, 25, 27, 30, 61-70).

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: IMMULOGIC PHARMACEUTICAL CORPORATION
(B) STREET: 610 LINCOLN STREET
(C) CITY: WALTHAM
(D) STATE: MASSACHUSETTS
(E) COUNTRY: USA
(F) POSTAL CODE (ZIP): 02154
(G) TELEPHONE: (617) 466-6000
(H) TELEFAX: (617) 466-6010

(ii) TITLE OF INVENTION: T CELL EPITOPES OF RYEGRASS POLLEN ALLERGENS

(iii) NUMBER OF SEQUENCES: 98

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: LAHIVE & COCKFIELD
(B) STREET: 60 State Street, suite # 510
(C) CITY: Boston
(D) STATE: Massachusetts
(E) COUNTRY: US
(F) ZIP: 02109-1875

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: ASCII text

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/106,016
(B) FILING DATE: 31-AUG-1993

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/031,001
(B) FILING DATE: 12-MAR-1993

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Amy E. Mandragouras
(B) REGISTRATION NUMBER: 36,207
(C) REFERENCE/DOCKET NUMBER: (IMI-040PC)

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 227-7400
(B) TELEFAX: (617) 227-5941

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1124 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 16..804

(ix) FEATURE:
(A) NAME/KEY: mat_peptide
(B) LOCATION: 85..804

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CAAATTCAAG ACAAG ATG GCG TCC TCC TCG TCG GTG CTC CTG GTG GTG GCG
51
                 Met Ala Ser Ser Ser Ser Val Leu Leu Val Val Ala
                 -23         -20                 -15

CTG TTC GCC GTG TTC CTG GGC AGC GCG CAT GGC ATC GCG AAG GTA CCA
99
Leu Phe Ala Val Phe Leu Gly Ser Ala His Gly Ile Ala Lys Val Pro
-10                 -5                  1                   5

CCG GGC CCC AAC ATC ACG GCC GAG TAC GGC GAC AAG TGG CTG GAC GCG
147
Pro Gly Pro Asn Ile Thr Ala Glu Tyr Gly Asp Lys Trp Leu Asp Ala
                10                  15                  20

AAG AGC ACC TGG TAT GGC AAG CCG ACC GGC GCC GGT CCC AAG GAC AAC
195
Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Ala Gly Pro Lys Asp Asn
            25                  30                  35

GGC GGC GCG TGC GGG TAC AAG GAC GTT GAC AAG GCG CCG TTC AAC GGC
243
Gly Gly Ala Cys Gly Tyr Lys Asp Val Asp Lys Ala Pro Phe Asn Gly
        40                  45                  50

ATG ACC GGC TGC GGC AAC ACC CCC ATC TTC AAG GAC GGC CGT GGC TGC
291
Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys Asp Gly Arg Gly Cys
    55                  60                  65

GGC TCC TGC TTC GAG ATC AAG TGC ACC AAG CCC GAG TCC TGC TCC GGC
339
Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro Glu Ser Cys Ser Gly
70                  75                  80                  85

GAG GCT GTC ACC GTC ACA ATC ACC GAC GAC AAC GAG GAG CCC ATC GCA
387
Glu Ala Val Thr Val Thr Ile Thr Asp Asp Asn Glu Glu Pro Ile Ala
                90                  95                  100

CCC TAC CAT TTC GAC CTC TCG GGC CAC GCG TTC GGG TCC ATG GCG AAG
435
Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe Gly Ser Met Ala Lys
            105                 110                 115

AAG GGC GAG GAG CAG AAG CTC CGC AGC GCC GGC GAG CTG GAG CTC CAG
483
Lys Gly Glu Glu Gln Lys Leu Arg Ser Ala Gly Glu Leu Glu Leu Gln
        120                 125                 130
```

2157596

```
TTC AGG CGG GTC AAG TGC AAG TAC CCG GAC GGC ACC AAG CCG ACA TTC
531
Phe Arg Arg Val Lys Cys Lys Tyr Pro Asp Gly Thr Lys Pro Thr Phe
    135                 140                 145

CAC GTC GAG AAG GCT TCC AAC CCC AAC TAC CTC GCT ATT CTG GTG AAG
579
His Val Glu Lys Ala Ser Asn Pro Asn Tyr Leu Ala Ile Leu Val Lys
150                 155                 160                 165

TAC GTC GAC GGC GAC GGT GAC GTG GTG GCG GTG GAC ATC AAG GAG AAG
627
Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val Asp Ile Lys Glu Lys
                170                 175                 180

GGC AAG GAT AAG TGG ATC GAG CTC AAG GAG TCG TGG GGA GCA GTC TGG
675
Gly Lys Asp Lys Trp Ile Glu Leu Lys Glu Ser Trp Gly Ala Val Trp
            185                 190                 195

AGG ATC GAC ACC CCC GAT AAG CTG ACG GGC CCA TTC ACC GTC CGC TAC
723
Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro Phe Thr Val Arg Tyr
        200                 205                 210

ACC ACC GAG GGC GGC ACC AAA TCC GAA GTC GAG GAT GTC ATC CCT GAG
771
Thr Thr Glu Gly Gly Thr Lys Ser Glu Val Glu Asp Val Ile Pro Glu
    215                 220                 225

GGC TGG AAG GCC GAC ACC TCC TAC TCG GCC AAG TGAGCAAGAA GTGGAGTGAT
824
Gly Trp Lys Ala Asp Thr Ser Tyr Ser Ala Lys
230                 235                 240

CTTCTTCCAA TCAGCTTAAT TTTGACTCAA GATCTCAAAT AATCCAGCCG CACATATATA
884

CGAGGCGGTG AGACATACAA GCTCCTCCAT GAGTATATTC ATTCATGCCG TATAGAGAGG
944

AGAAAGATGC CTGAATAAGA GTTTGAGGTC GACACCTTGT GAGAAGTGTA TATAGGAGGA
1004

ACCCAATCTG GCTCCATCTT TCTTTGCTCG CACGGTGTAC TGCTAAGGTT ATCTTCTAAC
1064

AGGCCAGATT AACCTACTAT CTAATATATG CAACGTATGG TCATTTTCCC TAAAAAAAAA
1124
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 263 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Ser Ser Ser Ser Val Leu Leu Val Val Ala Leu Phe Ala Val

```
-23         -20                 -15                 -10
Phe Leu Gly Ser Ala His Gly Ile Ala Lys Val Pro Pro Gly Pro Asn
        -5                  1                   5
Ile Thr Ala Glu Tyr Gly Asp Lys Trp Leu Asp Ala Lys Ser Thr Trp
 10                 15                  20                  25
Tyr Gly Lys Pro Thr Gly Ala Gly Pro Lys Asp Asn Gly Gly Ala Cys
             30                  35                  40
Gly Tyr Lys Asp Val Asp Lys Ala Pro Phe Asn Gly Met Thr Gly Cys
         45                  50                  55
Gly Asn Thr Pro Ile Phe Lys Asp Gly Arg Gly Cys Gly Ser Cys Phe
     60                  65                  70
Glu Ile Lys Cys Thr Lys Pro Glu Ser Cys Ser Gly Glu Ala Val Thr
 75                  80                  85
Val Thr Ile Thr Asp Asp Asn Glu Glu Pro Ile Ala Pro Tyr His Phe
 90                 95                  100                 105
Asp Leu Ser Gly His Ala Phe Gly Ser Met Ala Lys Lys Gly Glu Glu
            110                 115                 120
Gln Lys Leu Arg Ser Ala Gly Glu Leu Glu Leu Gln Phe Arg Arg Val
        125                 130                 135
Lys Cys Lys Tyr Pro Asp Gly Thr Lys Pro Thr Phe His Val Glu Lys
    140                 145                 150
Ala Ser Asn Pro Asn Tyr Leu Ala Ile Leu Val Lys Tyr Val Asp Gly
155                 160                 165
Asp Gly Asp Val Val Ala Val Asp Ile Lys Glu Lys Gly Lys Asp Lys
170                 175                 180                 185
Trp Ile Glu Leu Lys Glu Ser Trp Gly Ala Val Trp Arg Ile Asp Thr
            190                 195                 200
Pro Asp Lys Leu Thr Gly Pro Phe Thr Val Arg Tyr Thr Thr Glu Gly
        205                 210                 215
Gly Thr Lys Ser Glu Val Glu Asp Val Ile Pro Glu Gly Trp Lys Ala
    220                 225                 230
Asp Thr Ser Tyr Ser Ala Lys
235                 240
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ile Ala Lys Val Pro Pro Gly Pro Asn Ile Thr Ala Glu Tyr Gly Asp
1 5 10 15

Lys Trp Leu Asp
20

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ile Ala Lys Val Xaa Pro Gly Xaa Asn Ile Thr Ala Glu Tyr Gly Asp
1 5 10 15

Lys Trp Leu Asp
20

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Thr Ala Glu Tyr Gly Asp Lys Trp Leu Asp Ala Lys Ser Thr Trp Tyr
1 5 10 15

Gly Lys Pro Thr
20

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ala Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Ala Gly Pro Lys Asp
1 5 10 15

Asn Gly Gly Ala
20

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Gly Ala Gly Pro Lys Asp Asn Gly Gly Ala Cys Gly Tyr Lys Asn Val
1 5 10 15

Asp Lys Ala Pro
20

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Gly Ala Gly Pro Lys Asp Asn Gly Gly Ala Cys Gly Tyr Lys Asp Val
1 5 10 15

Asp Lys Ala Pro
20

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Cys Gly Tyr Lys Asp Val Asp Lys Ala Pro Phe Asn Gly Met Thr Gly
1 5 10 15

Cys Gly Asn Thr
20

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Phe Asn Gly Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys Asp Gly
1 5 10 15

Arg Gly Cys Gly
20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Pro Ile Phe Lys Asp Gly Arg Gly Cys Gly Ser Cys Phe Glu Ile Lys
1 5 10 15

Cys Thr Lys Pro
20

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro Glu Ser Cys Ser Gly Glu
1 5 10 15

Ala Val Thr Val
20

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Glu Ser Cys Ser Gly Glu Ala Val Thr Val Thr Ile Thr Asp Asp Asn
1 5 10 15

Glu Glu Pro Ile

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Thr Ile Thr Asp Asp Asn Glu Glu Pro Ile Ala Pro Tyr His Phe Asp
1 5 10 15

Leu Ser Gly His
20

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Ala Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe Gly Ser Met Ala
1 5 10 15

Asp Asp Gly Glu
20

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Ala Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe Gly Ser Met Ala
1 5 10 15

Lys Lys Gly Glu
20

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Ala Phe Gly Ser Met Ala Asp Asp Gly Glu Glu Gln Lys Leu Arg Ser
1 5 10 15

Ala Gly Glu Leu
20

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Ala Phe Gly Ser Met Ala Lys Lys Gly Glu Glu Gln Lys Leu Arg Ser
1 5 10 15

Ala Gly Glu Leu
20

(2) INFORMATION FOR SEQ ID NO:19:

2157596

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Glu Gln Lys Leu Arg Ser Ala Gly Glu Leu Glu Leu Gln Phe Arg Arg
1 5 10 15

Val Lys Cys Lys
20

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Glu Leu Gln Phe Arg Arg Val Lys Cys Lys Tyr Pro Asp Asp Thr Lys
1 5 10 15

Pro Thr Phe His
20

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Tyr Pro Asp Asp Thr Lys Pro Thr Phe His Val Glu Lys Ala Ser Asn
1 5 10 15

Pro Asn Tyr Leu
20

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Val Glu Lys Ala Ser Asn Pro Asn Tyr Leu Ala Ile Leu Val Lys Tyr
1 5 10 15

Val Asp Gly Asp
20

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu Ala Ile Leu Val Lys Tyr
1 5 10 15

Val Asp Gly Asp
20

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Ala Ile Leu Val Lys Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val
1 5 10 15

Asp Ile Lys Glu
20

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Gly Asp Val Val Ala Val Asp Ile Lys Glu Lys Gly Lys Asp Lys Trp
1 5 10 15

Ile Glu Leu Lys
20

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Lys Gly Lys Asp Lys Trp Ile Glu Leu Lys Glu Ser Trp Gly Ala Val
1 5 10 15

Trp Arg Ile Asp
20

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Glu Ser Trp Gly Ala Val Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr
1 5 10 15

Gly Pro Phe Thr
20

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Thr Pro Asp Lys Leu Thr Gly Pro Phe Thr Val Arg Tyr Thr Thr Glu
1 5 10 15

Gly Gly Thr Lys
20

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Val Arg Tyr Thr Thr Glu Gly Gly Thr Lys Ser Glu Val Glu Asp Val
1 5 10 15

Ile Pro Glu Gly
20

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Ser Glu Val Glu Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr Ser
1 5 10 15

Tyr Ser Ala Lys
20

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Asp Glu Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu Ala Ile Leu Val
1 5 10 15

Lys Tyr Val Asp Gly Asp
20

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Asp Glu Ala Glu Lys Gly Ser Asn Pro Asn Tyr Leu Ala Ile Leu Val
1 5 10 15

Lys Tyr Val Asp Gly Asp
20

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Lys Lys Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu Ala Ile Leu Val
1 5 10 15

Lys Lys

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu Ala Ile Leu Asp Glu
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Ala Glu Lys Gly Ser Asn Pro Asn Tyr Leu Ala Ile Leu Asp Glu
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

Asp Glu Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu Ala Ile Asp Glu
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Lys Lys Ala Glu Lys Gly Ser Asn Pro Asn Tyr Leu Ala Ile Leu Val
1 5 10 15

Lys Lys

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Asp Glu Pro Asn Tyr Leu Ala Ile Leu Val Lys Tyr Val Asp Glu
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

Gly Asp Val Val Ala Val Asp Ile Lys Glu Lys Gly Lys Asp Lys
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Val Ala Val Asp Ile Lys Glu Lys Gly Lys Asp Lys Trp Ile Glu
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Ala Val Asp Ile Lys Glu Lys Gly Lys Asp Lys Trp Ile Glu Leu
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Asp Ile Lys Glu Lys Gly Lys Asp Lys Trp Ile Glu Leu Lys
1 5 10

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

Trp Gly Ala Val Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

Gly Ala Val Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly
1 5 10

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro Phe Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

Glu Ser Trp Gly Ala Val Trp Arg Ile Asp Thr Pro Asp Lys
1 5 10

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Ala Gly Ala Val Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

Ser Glu Val Glu Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

Glu Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr Ser Tyr Ser
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Ile Pro Glu Gly Trp Lys Ala Asp Thr Ser Tyr Ser Ala Lys
1 5 10

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 723 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..720

2157596

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

```
ATC CCG AAG GTG CCC CCG GGC CCG AAC ATC ACG GCG ACC TAC GGT GAC
48
Ile Pro Lys Val Pro Pro Gly Pro Asn Ile Thr Ala Thr Tyr Gly Asp
 1               5                  10                  15

AAG TGG CTG GAC GCG AAG AGC ACA TGG TAC GGC AAG CCG ACG GGC GCC
96
Lys Trp Leu Asp Ala Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Ala
            20                  25                  30

GGC CCC AAG GAC AAC GGC GGC GCG TGC GGG TAC AAG GAC GTG GAC AAG
144
Gly Pro Lys Asp Asn Gly Gly Ala Cys Gly Tyr Lys Asp Val Asp Lys
        35                  40                  45

GCG CCG TTC AAC GGC ATG ACC GGG TGC GGC AAC ACC CCC ATC TTC AAG
192
Ala Pro Phe Asn Gly Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys
    50                  55                  60

GAC GGG CGC GGG TGC GGT TCC TGC TTC GAG ATC AAG TGC ACG AAG CCC
240
Asp Gly Arg Gly Cys Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro
65                  70                  75                  80

GAG TCG TGC TCC GGC GAG GCC GTC ACC GTC CAC ATC ACC GAC GAC AAC
288
Glu Ser Cys Ser Gly Glu Ala Val Thr Val His Ile Thr Asp Asp Asn
                85                  90                  95

GAG GAG CCC ATC GCG CCC TAC CAC TTC GAC CTT TCC GGC CAC GCG TTC
336
Glu Glu Pro Ile Ala Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe
            100                 105                 110

GGT TCC ATG GCG AAG AAG GGC GAG GAG CAG AAG CTG CGC AGC GCG GGC
384
Gly Ser Met Ala Lys Lys Gly Glu Glu Gln Lys Leu Arg Ser Ala Gly
        115                 120                 125

GAG CTG GAG CTG CAG TTT AGG CGG GTG AAG TGC AAG TAC CCC GAG GGC
432
Glu Leu Glu Leu Gln Phe Arg Arg Val Lys Cys Lys Tyr Pro Glu Gly
    130                 135                 140

ACC AAG GTG ACC TTC CAC GTC GAG AAG GGT TCC AAC CCC AAC TAC CTG
480
Thr Lys Val Thr Phe His Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu
145                 150                 155                 160

GCG CTG CTG GTG AAG TAC GTC GAC GGC GAC GGC GAC GTG GTG GCG GTG
528
Ala Leu Leu Val Lys Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val
                165                 170                 175

GAT ATC AAG GAG AAG GGC AAG GAC AAG TGG ATC GCG CTC AAG GAG TCA
576
Asp Ile Lys Glu Lys Gly Lys Asp Lys Trp Ile Ala Leu Lys Glu Ser
            180                 185                 190

TGG GGA GCC ATC TGG AGG GTG GAC ACC CCC GAC AAG CTG ACG GGC CCA
624
```

Trp Gly Ala Ile Trp Arg Val Asp Thr Pro Asp Lys Leu Thr Gly Pro
195 200 205

```
TTC ACC GTT CGC TAC ACC ACC GAG GGA GGC ACC AAG TCC GAA GTT GAG
672
Phe Thr Val Arg Tyr Thr Thr Glu Gly Gly Thr Lys Ser Glu Val Glu
    210                 215                 220

GAC GTC ATC CCC GAG GGC TGG AAG GCC GAC GCC AGC TAC GAG TCC AAG
720
Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Ala Ser Tyr Glu Ser Lys
225                 230                 235                 240

TGA
723
```

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

```
Ile Pro Lys Val Pro Pro Gly Pro Asn Ile Thr Ala Thr Tyr Gly Asp
 1               5                  10                  15

Lys Trp Leu Asp Ala Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Ala
            20                  25                  30

Gly Pro Lys Asp Asn Gly Gly Ala Cys Gly Tyr Lys Asp Val Asp Lys
        35                  40                  45

Ala Pro Phe Asn Gly Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys
    50                  55                  60

Asp Gly Arg Gly Cys Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro
65                  70                  75                  80

Glu Ser Cys Ser Gly Glu Ala Val Thr Val His Ile Thr Asp Asp Asn
                85                  90                  95

Glu Glu Pro Ile Ala Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe
            100                 105                 110

Gly Ser Met Ala Lys Lys Gly Glu Glu Gln Lys Leu Arg Ser Ala Gly
        115                 120                 125

Glu Leu Glu Leu Gln Phe Arg Arg Val Lys Cys Lys Tyr Pro Glu Gly
    130                 135                 140

Thr Lys Val Thr Phe His Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu
145                 150                 155                 160

Ala Leu Leu Val Lys Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val
                165                 170                 175

Asp Ile Lys Glu Lys Gly Lys Asp Lys Trp Ile Ala Leu Lys Glu Ser
            180                 185                 190

Trp Gly Ala Ile Trp Arg Val Asp Thr Pro Asp Lys Leu Thr Gly Pro
        195                 200                 205
```

Phe Thr Val Arg Tyr Thr Thr Glu Gly Gly Thr Lys Ser Glu Val Glu
210 215 220

Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Ala Ser Tyr Glu Ser Lys
225 230 235 240

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 723 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..720

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "Xaa is Ser, Pro or Ala"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

ATC GCG AAG GTT CCC CCC GGC CCG AAC ATC ACG GCG ACC TAC GGC GAC
48
Ile Ala Lys Val Pro Pro Gly Pro Asn Ile Thr Ala Thr Tyr Gly Asp
1 5 10 15

AAG TGG CTT GAC GCG AAG AGC ACC TGG TAC GGC AAG CCG ACC GGN BCC
96
Lys Trp Leu Asp Ala Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Xaa
20 25 30

GGT CCC AAG GAC AAC GGC GGC GCG TGC GGA TAC AAG GAC GTG GAC AAG
144
Gly Pro Lys Asp Asn Gly Gly Ala Cys Gly Tyr Lys Asp Val Asp Lys
35 40 45

CCC CCG TTC AGC GGC ATG ACC GGC TGC GGC AAC ACC CCC ATC TTC AAG
192
Pro Pro Phe Ser Gly Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys
50 55 60

TCC GGC CGC GGC TGC GGC TCC TGC TTC GAG ATC AAG TGC ACC AAG CCC
240
Ser Gly Arg Gly Cys Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro
65 70 75 80

GAG TCC TGC TCC GGG GAG CCC GTC CTG GTC CAC ATC ACC GAC GAC AAC
288
Glu Ser Cys Ser Gly Glu Pro Val Leu Val His Ile Thr Asp Asp Asn
85 90 95

GAG GAG CCC ATC GCC GCC TAC CAC TTC GAC CTC TCC GGC AAG GCG TTC
336
Glu Glu Pro Ile Ala Ala Tyr His Phe Asp Leu Ser Gly Lys Ala Phe
100 105 110

```
GGG GCC ATG GCC AAG AAG GGT GAG GAG CAG AAG CTG CGC AGC GCC GGC
384
Gly Ala Met Ala Lys Lys Gly Glu Glu Gln Lys Leu Arg Ser Ala Gly
        115                 120                 125

GAG CTG GAG CTC AAG TTC CGC CGC GTC AAG TGC GAG TAC CCG AAG GGC
432
Glu Leu Glu Leu Lys Phe Arg Arg Val Lys Cys Glu Tyr Pro Lys Gly
    130                 135                 140

ACC AAG GTT ACC TTC CAC GTC GAG AAG GGG TCC AAC CCC AAC TAC CTT
480
Thr Lys Val Thr Phe His Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu
145                 150                 155                 160

GCG CTG CTG GTG AAG TAC GTC GAC GGC GAC GGG GAC GTG GTG GCG GTG
528
Ala Leu Leu Val Lys Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val
                165                 170                 175

GAC ATC AAG CAG AAG GGC AAG GAC AAG TGG ATC GAG CTC AAG GAG TCG
576
Asp Ile Lys Gln Lys Gly Lys Asp Lys Trp Ile Glu Leu Lys Glu Ser
            180                 185                 190

TGG GGA GCC GTC TGG AGG ATC GAC ACC CCC GAC AAG CTC ACC GGC CCC
624
Trp Gly Ala Val Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro
        195                 200                 205

TTC ACC GTC CGC TAC ACC ACC GAG GGC GGC ACC AAG GCC GAA GCC GAG
672
Phe Thr Val Arg Tyr Thr Thr Glu Gly Gly Thr Lys Ala Glu Ala Glu
    210                 215                 220

GAC GTC ATC CCC GAG GGC TGG AAG GCC GAC ACC GCC TAC GAG GCC AAG
720
Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr Ala Tyr Glu Ala Lys
225                 230                 235                 240

TGA
723
```

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 32
(D) OTHER INFORMATION: /note= "Xaa is Ser, Pro or Ala"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Ile Ala Lys Val Pro Pro Gly Pro Asn Ile Thr Ala Thr Tyr Gly Asp
1 5 10 15

Lys Trp Leu Asp Ala Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Xaa
20 25 30

Gly Pro Lys Asp Asn Gly Gly Ala Cys Gly Tyr Lys Asp Val Asp Lys
35 40 45

Pro Pro Phe Ser Gly Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys
50 55 60

Ser Gly Arg Gly Cys Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro
65 70 75 80

Glu Ser Cys Ser Gly Glu Pro Val Leu Val His Ile Thr Asp Asp Asn
85 90 95

Glu Glu Pro Ile Ala Ala Tyr His Phe Asp Leu Ser Gly Lys Ala Phe
100 105 110

Gly Ala Met Ala Lys Lys Gly Glu Glu Gln Lys Leu Arg Ser Ala Gly
115 120 125

Glu Leu Glu Leu Lys Phe Arg Arg Val Lys Cys Glu Tyr Pro Lys Gly
130 135 140

Thr Lys Val Thr Phe His Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu
145 150 155 160

Ala Leu Leu Val Lys Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val
165 170 175

Asp Ile Lys Gln Lys Gly Lys Asp Lys Trp Ile Glu Leu Lys Glu Ser
180 185 190

Trp Gly Ala Val Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro
195 200 205

Phe Thr Val Arg Tyr Thr Thr Glu Gly Gly Thr Lys Ala Glu Ala Glu
210 215 220

Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr Ala Tyr Glu Ala Lys
225 230 235 240

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 723 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..720

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

```
ATC GCG AAG GTG CCC CCG GGT CCG AAC ATC ACG GCG ACC TAC GGC GAC
48
Ile Ala Lys Val Pro Pro Gly Pro Asn Ile Thr Ala Thr Tyr Gly Asp
 1               5                  10                  15

AAG TGG CTC GAC GCG AAG AGC ACA TGG TAC GGC AAG CCG ACG GGG GCC
96
Lys Trp Leu Asp Ala Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Ala
            20                  25                  30

GGT CCC AAG GAC AAC GGC GGC GCT TGC GGG TAC AAG GAC GTG GAC AAG
144
Gly Pro Lys Asp Asn Gly Gly Ala Cys Gly Tyr Lys Asp Val Asp Lys
        35                  40                  45

CCC CCG TTC AGC GGC ATG ACC GGC TGC GGC AAC ACC CCC ATC TTC AAG
192
Pro Pro Phe Ser Gly Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys
    50                  55                  60

TCC GGC CGT GGC TGC GGC TCC TGC TTT GAG ATC AAG TGC ACG AAG CCC
240
Ser Gly Arg Gly Cys Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro
65                  70                  75                  80

GAG GCC TGC TCC GGC GAG CCC GTG GTA GTC CAC ATC ACC GAC GAC AAC
288
Glu Ala Cys Ser Gly Glu Pro Val Val Val His Ile Thr Asp Asp Asn
                85                  90                  95

GAG GAG CCC ATC GCC CCC TAC CAC TTC GAC CTC TCC GGC CAC GCG TTC
336
Glu Glu Pro Ile Ala Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe
            100                 105                 110

GGG GCG ATG GCC AAG AAG GGC GAT GAG CAG AAG CTG CGC ACG GCC GGC
384
Gly Ala Met Ala Lys Lys Gly Asp Glu Gln Lys Leu Arg Thr Ala Gly
        115                 120                 125

GAG CTG GAG CTC CAG TTC CGG CGC GTC AAG TGC AAG TAC CCG GAG GGG
432
Glu Leu Glu Leu Gln Phe Arg Arg Val Lys Cys Lys Tyr Pro Glu Gly
    130                 135                 140
```

```
ACC AAG GTG ACC TTC CAC GTG GAG AAG GGG TCC AAC CCC AAC TAC CTG
480
Thr Lys Val Thr Phe His Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu
145                 150                 155                 160

GCG CTG CTT GTG AAG TAC GTT AAC GGC GAC GGA GAC GTG GTG GCG GTG
528
Ala Leu Leu Val Lys Tyr Val Asn Gly Asp Gly Asp Val Val Ala Val
                165                 170                 175

GAC ATC AAG GAG AAG GGC AAG GAC AAG TGG ATC GAG CTC AAG GAG TCG
576
Asp Ile Lys Glu Lys Gly Lys Asp Lys Trp Ile Glu Leu Lys Glu Ser
            180                 185                 190

TGG GGA GCC ATC TGG AGG ATC GAC ACT CCC GAC AAG CTC ACG GGC CCC
624
Trp Gly Ala Ile Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro
        195                 200                 205

TTC ACC GTC CGC TAC ACC ACC GAG GGC GGC ACC AAG ACC GAA GCC GAG
672
Phe Thr Val Arg Tyr Thr Thr Glu Gly Gly Thr Lys Thr Glu Ala Glu
    210                 215                 220

GAC GTC ATC CCT GAG GGC TGG AAG GCC GAC ACC AGC TAC GAG TCC AAG
720
Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr Ser Tyr Glu Ser Lys
225                 230                 235                 240

TGA
723
```

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

```
Ile Ala Lys Val Pro Pro Gly Pro Asn Ile Thr Ala Thr Tyr Gly Asp
 1               5                  10                  15

Lys Trp Leu Asp Ala Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Ala
            20                  25                  30

Gly Pro Lys Asp Asn Gly Gly Ala Cys Gly Tyr Lys Asp Val Asp Lys
        35                  40                  45

Pro Pro Phe Ser Gly Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys
    50                  55                  60

Ser Gly Arg Gly Cys Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro
65                  70                  75                  80

Glu Ala Cys Ser Gly Glu Pro Val Val Val His Ile Thr Asp Asp Asn
                85                  90                  95

Glu Glu Pro Ile Ala Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe
```

100 105

Gly Ala Met Ala Lys Lys Gly Asp Glu Gln Lys Leu Arg Thr Ala Gly
115 120 125

Glu Leu Glu Leu Gln Phe Arg Arg Val Lys Cys Lys Tyr Pro Glu Gly
130 135 140

Thr Lys Val Thr Phe His Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu
145 150 155 160

Ala Leu Leu Val Lys Tyr Val Asn Gly Asp Gly Asp Val Val Ala Val
165 170 175

Asp Ile Lys Glu Lys Gly Lys Asp Lys Trp Ile Glu Leu Lys Glu Ser
180 185 190

Trp Gly Ala Ile Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro
195 200 205

Phe Thr Val Arg Tyr Thr Thr Glu Gly Gly Thr Lys Thr Glu Ala Glu
210 215 220

Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr Ser Tyr Glu Ser Lys
225 230 235 240

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION: 45
(D) OTHER INFORMATION: /note= "Xaa is Asn or Asp"

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION: 144
(D) OTHER INFORMATION: /note= "Xaa is Asp or Gly"

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION: 154
(D) OTHER INFORMATION: /note= "Xaa is Gly or Ala"

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION: 187
(D) OTHER INFORMATION: /note= "Xaa is Ile or Thr"

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION: 223
(D) OTHER INFORMATION: /note= "Xaa is Val or Phe"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Ile Ala Lys Val Pro Pro Gly Pro Asn Ile Thr Ala Glu Tyr Gly Asp
1 5 10 15

Lys Trp Leu Asp Ala Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Ala
20 25 30

Gly Pro Lys Asp Asn Gly Gly Ala Cys Gly Tyr Lys Xaa Val Asp Lys
35 40 45

Ala Pro Phe Asn Gly Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys
50 55 60

Asp Gly Arg Gly Cys Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro
65 70 75 80

Glu Ser Cys Ser Gly Glu Ala Val Thr Val Thr Ile Thr Asp Asp Asn
85 90 95

Glu Glu Pro Ile Ala Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe
100 105 110

Gly Ser Met Ala Lys Lys Gly Glu Glu Gln Lys Leu Arg Ser Ala Gly
115 120 125

Glu Leu Glu Leu Gln Phe Arg Arg Val Lys Cys Lys Tyr Pro Asp Xaa
130 135 140

Thr Lys Pro Thr Phe His Val Glu Lys Xaa Ser Asn Pro Asn Tyr Leu
145 150 155 160

Ala Ile Leu Val Lys Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val
165 170 175

Asp Ile Lys Glu Lys Gly Lys Asp Lys Trp Xaa Glu Leu Lys Glu Ser
180 185 190

Trp Gly Ala Val Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro
195 200 205

Phe Thr Val Arg Tyr Thr Thr Glu Gly Gly Thr Lys Ser Glu Xaa Glu
210 215 220

Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr Ser Tyr Ser Ala Lys
225 230 235 240

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY:

(B) LOCATION: 199
(D) OTHER INFORMATION: /note= "Xaa is Val or Ile"

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION: 235
(D) OTHER INFORMATION: /note= "Xaa is Ala or Thr"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

Ile Pro Lys Val Pro Pro Gly Pro Asn Ile Thr Ala Thr Tyr Gly Asp
1 5 10 15

Lys Trp Leu Asp Ala Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Ala
20 25 30

Gly Pro Lys Asp Asn Gly Gly Ala Cys Gly Tyr Lys Asp Val Asp Lys
35 40 45

Ala Pro Phe Asn Gly Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys
50 55 60

Asp Gly Arg Gly Cys Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro
65 70 75 80

Glu Ser Cys Ser Gly Glu Ala Val Thr Val His Ile Thr Asp Asp Asn
85 90 95

Glu Glu Pro Ile Ala Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe
100 105 110

Gly Ser Met Ala Lys Lys Gly Glu Glu Gln Lys Leu Arg Ser Ala Gly
115 120 125

Glu Leu Glu Leu Gln Phe Arg Arg Val Lys Cys Lys Tyr Pro Glu Gly
130 135 140

Thr Lys Val Thr Phe His Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu
145 150 155 160

Ala Leu Leu Val Lys Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val
165 170 175

Asp Ile Lys Glu Lys Gly Lys Asp Lys Trp Ile Ala Leu Lys Glu Ser
180 185 190

Trp Gly Ala Ile Trp Arg Xaa Asp Thr Pro Asp Lys Leu Thr Gly Pro
195 200 205

Phe Thr Val Arg Tyr Thr Thr Glu Gly Gly Thr Lys Ser Glu Val Glu
210 215 220

Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Xaa Ser Tyr Glu Ser Lys
225 230 235 240

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

Ile Ala Lys Val Pro Pro Gly Pro Asn Ile Thr Ala Thr Tyr Gly Asp
1 5 10 15

Lys Trp Leu Asp Ala Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Ala
20 25 30

Gly Pro Lys Asp Asn Gly Gly Ala Cys Gly Tyr Lys Asp Val Asp Lys
35 40 45

Pro Pro Phe Ser Gly Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys
50 55 60

Ser Gly Arg Gly Cys Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro
65 70 75 80

Glu Ala Cys Ser Gly Glu Pro Val Val Val His Ile Thr Asp Asp Asn
85 90 95

Glu Glu Pro Ile Ala Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe
100 105 110

Gly Ala Met Ala Lys Lys Gly Asp Glu Gln Lys Leu Arg Thr Ala Gly
115 120 125

Glu Leu Glu Leu Gln Phe Arg Arg Val Lys Cys Lys Tyr Pro Glu Gly
130 135 140

Thr Lys Val Thr Phe His Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu
145 150 155 160

Ala Leu Leu Val Lys Tyr Val Asn Gly Asp Gly Asp Val Val Ala Val
165 170 175

Asp Ile Lys Glu Lys Gly Lys Asp Lys Trp Ile Glu Leu Lys Glu Ser
180 185 190

Trp Gly Ala Ile Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro
195 200 205

Phe Thr Val Arg Tyr Thr Thr Glu Gly Gly Thr Lys Thr Glu Ala Glu
210 215 220

Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr Ser Tyr Glu Ser Lys
225 230 235 240

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION: 88
(D) OTHER INFORMATION: /note= "Xaa is Val or Ile"

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION: 90
(D) OTHER INFORMATION: /note= "Xaa is Val or Ile"

(ix) FEATURE:
(A) NAME/KEY:
(B) LOCATION: 180
(D) OTHER INFORMATION: /note= "Xaa is Gln or Glu"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

Ile Ala Lys Val Pro Pro Gly Pro Asn Ile Thr Ala Thr Tyr Gly Asp
1 5 10 15

Lys Trp Leu Asp Ala Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Xaa
20 25 30

Gly Pro Lys Asp Asn Gly Gly Ala Cys Gly Tyr Lys Asp Val Asp Lys
35 40 45

Pro Pro Phe Ser Gly Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys
50 55 60

Ser Gly Arg Gly Cys Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro
65 70 75 80

Glu Ser Cys Ser Gly Glu Pro Xaa Leu Xaa His Ile Thr Asp Asp Asn

85 90 95

Glu Glu Pro Ile Ala Ala Tyr His Phe Asp Leu Ser Gly Lys Ala Phe
100 105 110

Gly Ala Met Ala Lys Lys Gly Glu Glu Gln Lys Leu Arg Ser Ala Gly
115 120 125

Glu Leu Glu Leu Lys Phe Arg Arg Val Lys Cys Glu Tyr Pro Lys Gly
130 135 140

Thr Lys Val Thr Phe His Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu
145 150 155 160

Ala Leu Leu Val Lys Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val
165 170 175

Asp Ile Lys Xaa Lys Gly Lys Asp Lys Trp Ile Glu Leu Lys Glu Ser
180 185 190

Trp Gly Ala Val Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro
195 200 205

Phe Thr Val Arg Tyr Thr Thr Glu Gly Gly Thr Lys Ala Glu Ala Glu
210 215 220

Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr Ala Tyr Glu Ala Lys
225 230 235 240

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu Ala Leu Leu Val Lys Tyr
1 5 10 15

Val Asp Gly Asp
20

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

Val Glu Lys Gly Ser Asn Pro Asn Tyr Leu Ala Leu Leu Val Lys Tyr
1 5 10 15

Val Asn Gly Asp
20

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

Gly Asp Val Val Ala Val Asp Ile Lys Glu Lys Gly Lys Asp Lys Trp
1 5 10 15

Ile Ala Leu Lys
20

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Gly Asp Val Val Ala Val Asp Ile Lys Gln Lys Gly Lys Asp Lys Trp
1 5 10 15

Ile Glu Leu Lys
20

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Glu Ser Trp Gly Ala Ile Trp Arg Ile Asp Thr Pro Asp Lys Leu Thr
1 5 10 15

Gly Pro Phe Thr
20

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

Glu Ser Trp Gly Ala Ile Trp Arg Val Asp Thr Pro Asp Lys Leu Thr
1 5 10 15

Gly Pro Phe Thr
20

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Thr Glu Ala Glu Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr Ser
1 5 10 15

Tyr Glu Ser Lys
20

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

Ala Glu Ala Glu Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr Ala

1 5 10 15

Tyr Glu Ala Lys
20

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

Ser Glu Val Glu Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Ala Ser
1 5 10 15

Tyr Glu Ser Lys
20

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

Ser Glu Val Glu Asp Val Ile Pro Glu Gly Trp Lys Ala Asp Thr Ser
1 5 10 15

Tyr Glu Ser Lys
20

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

GGGTCTAGAG GTACCGTCCG ATCGATCATT
30

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

AATGATCGAT GCT
13

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

GGGTCTAGAG GTACCGTCCG
20

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

CCCTGCAGAT TATTTGAGAT CTTGAG
26

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

CCCTGCAGTC ATGCTCACTT GGCCGAGTA
29

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

GAGTACGGCG ACAAGTGGC
19

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

TTCGAGATCA AGTGCACC
18

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

GTGACAGCCT CGCCGG
16

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

GGGAATTCCA TGGCGAAGAA GGGC
24

(2) INFORMATION FOR SEQ ID NO:80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

GTGCCGTCCG GGTACT
16

(2) INFORMATION FOR SEQ ID NO:81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

CCGTCGACGT ACTTCA
16

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

GGAGTCGTGG GGAGCAGTC
19

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

GGGTCTAGAG GTACCGTCC
19

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

TTGGATCCTA CGGCAAGCCG ACCGGC
26

(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

TTGGATCCAT CCCGAAGGTG CCCCCGGG
28

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

AGGTGACCTT CCACGTCG
18

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

TTGGATCCTG GCGCTGCTGG TGAAGTA
27

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

TTGAATTCAT CCCGAAGGTG CCCCCG
26

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

TTGGTACCTC ACTTGGACTC GTAGCT
26

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

CCGAATTCGT GGAGAAGGGG TCCAA
25

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 22
(D) OTHER INFORMATION: /note= "Xaa is Iosine"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

TTAGGATCCT CACTTATCAT ANGACGTATC
30

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

TTGAATTCCT TGTCATTGCC CTTCTG
26

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

AAGAATTCCT TCTGCTTGAT GTCCAC
26

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

ATGAATTCGA GTCGTGGGGA GCCGTC
26

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

ATGAATTCGT CTGGAGGATC GACACC
26

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

ATGAATTCAT CGCAAAGGTT CCCCCC
26

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

TTTGGATCCT CACTTGGACT CGTAGCT
27

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

TTGAATTCTC GCGAAGGTGC CCCCG
25

```
                                                                    51
CAAATTCAAG ACAAG ATG GCG TCC TCC TCG TCG GTG CTC CTG GTG GTG GCG
                 Met Ala Ser Ser Ser Val Leu Leu Val Val Ala
                  -23      -20             -15
                                                                    99
CTG TTC GCC GTG TTC CTG GGC AGC GCG CAT GGC ATC GCG AAG GTA CCA
Leu Phe Ala Val Phe Leu Gly Ser Ala His Gly Ile Ala Lys Val Pro
     -10             -5                  1               5
                                                                   147
CCG GGC CCC AAC ATC ACG GCC GAG TAC GGC GAC AAG TGG CTG GAC GCG
Pro Gly Pro Asn Ile Thr Ala Glu Tyr Gly Asp Lys Trp Leu Asp Ala
         10                  15                  20
                                                                   195
AAG AGC ACC TGG TAT GGC AAG CCG ACC GGC GCC GGT CCC AAG GAC AAC
Lys Ser Thr Trp Tyr Gly Lys Pro Thr Gly Ala Gly Pro Lys Asp Asn
             25                  30                  35
                                                                   243
GGC GGC GCG TGC GGG TAC AAG GAC GTT GAC AAG GCG CCG TTC AAC GGC
Gly Gly Ala Cys Gly Tyr Lys Asp Val Asp Lys Ala Pro Phe Asn Gly
         40                  45                  50
                                                                   291
ATG ACC GGC TGC GGC AAC ACC CCC ATC TTC AAG GAC GGC CGT GGC TGC
Met Thr Gly Cys Gly Asn Thr Pro Ile Phe Lys Asp Gly Arg Gly Cys
     55                  60                  65
                                                                   339
GGC TCC TGC TTC GAG ATC AAG TGC ACC AAG CCC GAG TCC TGC TCC GGC
Gly Ser Cys Phe Glu Ile Lys Cys Thr Lys Pro Glu Ser Cys Ser Gly
 70                  75                  80                  85
```

Fig. 1

```
                                                                  387
GAG GCT GTC ACC GTC ACA ATC ACC GAC GAC AAC GAG GAG CCC ATC GCA
Glu Ala Val Thr Val Thr Ile Thr Asp Asp Asn Glu Glu Pro Ile Ala
          90                  95                  100

                                                                  435
CCC TAC CAT TTC GAC CTC TCG GGC CAC GCG TTC GGG TCC ATG GCG AAG
Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe Gly Ser Met Ala Lys
                  105                 110                 115

                                                                  483
AAG GGC GAG GAG CAG AAG CTC CGC AGC GCC GGC GAG CTG GAG CTC CAG
Lys Gly Glu Glu Gln Lys Leu Arg Ser Ala Gly Glu Leu Glu Leu Gln
                  120                 125                 130

                                                                  531
TTC AGG CGG GTC AAG TGC AAG TAC CCG GAC GGC ACC AAG CCG ACA TTC
Phe Arg Arg Val Lys Cys Lys Tyr Pro Asp Gly Thr Lys Pro Thr Phe
          135                 140                 145

                                                                  579
CAC GTC GAG AAG GCT TCC AAC CCC AAC TAC CTC GCT ATT CTG GTG AAG
His Val Glu Lys Ala Ser Asn Pro Asn Tyr Leu Ala Ile Leu Val Lys
  150                 155                 160                 165

                                                                  627
TAC GTC GAC GGC GAC GGT GAC GTG GTG GCG GTG GAC ATC AAG GAG AAG
Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val Asp Ile Lys Glu Lys
                  170                 175                 180
```

Fig. 1 cont.

```
GGC AAG GAT AAG TGG ATC GAG CTC AAG GAG TCG TGG GGA GCA GTC TGG  675
Gly Lys Asp Lys Trp Ile Glu Leu Lys Glu Ser Trp Gly Ala Val Trp
        185             190             195

AGG ATC GAC ACC CCC GAT AAG CTG ACG GGC CCA TTC ACC GTC CGC TAC  723
Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro Phe Thr Val Arg Tyr
    200             205             210

ACC ACC GAG GGC GGC ACC AAA TCC GAA GTC GAG GAT GTC ATC CCT GAG  771
Thr Thr Glu Gly Gly Thr Lys Ser Glu Val Glu Asp Val Ile Pro Glu
215             220             225

GGC TGG AAG GCC GAC ACC TCC TAC TCG GCC AAG TGAGCA               810
Gly Trp Lys Ala Asp Thr Ser Tyr Ser Ala Lys
230             235             240
```

Fig. 1 cont.

2157596

PEPTIDE NAME	PEPTIDE SEQUENCE
LPI-1	IAKVPPGPNITAEYGDKWLD
LPI-1.1	IAKVXPGXNITAEYGDKWLD
LPI-2	TAEYGDKWLDAKSTWYGKPT
LPI-3	AKSTWYGKPTGAGPKDNGGA
LPI-4	GAGPKDNGGACGYKNVDKAP
LPI-4.1	GAGPKDNGGACGYKDVDKAP
LPI-5	CGYKDVDKAPFNGMTGCGNT
LPI-6	FNGMTGCGNTPIFKDGRGCG
LPI-7	PIFKDGRGCGSCFEIKCTKP
LPI-8	SCFEIKCTKPESCSGEAVTV
LPI-9	ESCSGEAVTVTITDDNEEPI
LPI-10	TITDDNEEPIAPYHFDLSGH
LPI-11	APYHFDLSGHAFGSMADDGE
LPI-11.1	APYHFDLSGHAFGSMAKKGE
LPI-12	AFGSMADDGEEQKLRSAGEL
LPI-12.1	AFGSMAKKGEEQKLRSAGEL
LPI-13	EQKLRSAGELELQFRRVKCK
LPI-14	ELQFRRVKCKYPDDTKPTFH
LPI-15	YPDDTKPTFHVEKASNPNYL
LPI-16	VEKASNPNYLAILVKYVDGD
LPI-16.1	VEKGSNPNYLAILVKYVDGD
LPI-17	AILVKYVDGDGDVVAVDIKE
LPI-18	GDVVAVDIKEKGKDKWIELK
LPI-19	KGKDKWIELKESWGAVWRID
LPI-20	ESWGAVWRIDTPDKLTGPFT
LPI-21	TPDKLTGPFTVRYTTEGGTK
LPI-22	VRYTTEGGTKSEVEDVIPEG
LPI-23	SEVEDVIPEGWKADTSYSAK

Fig. 2

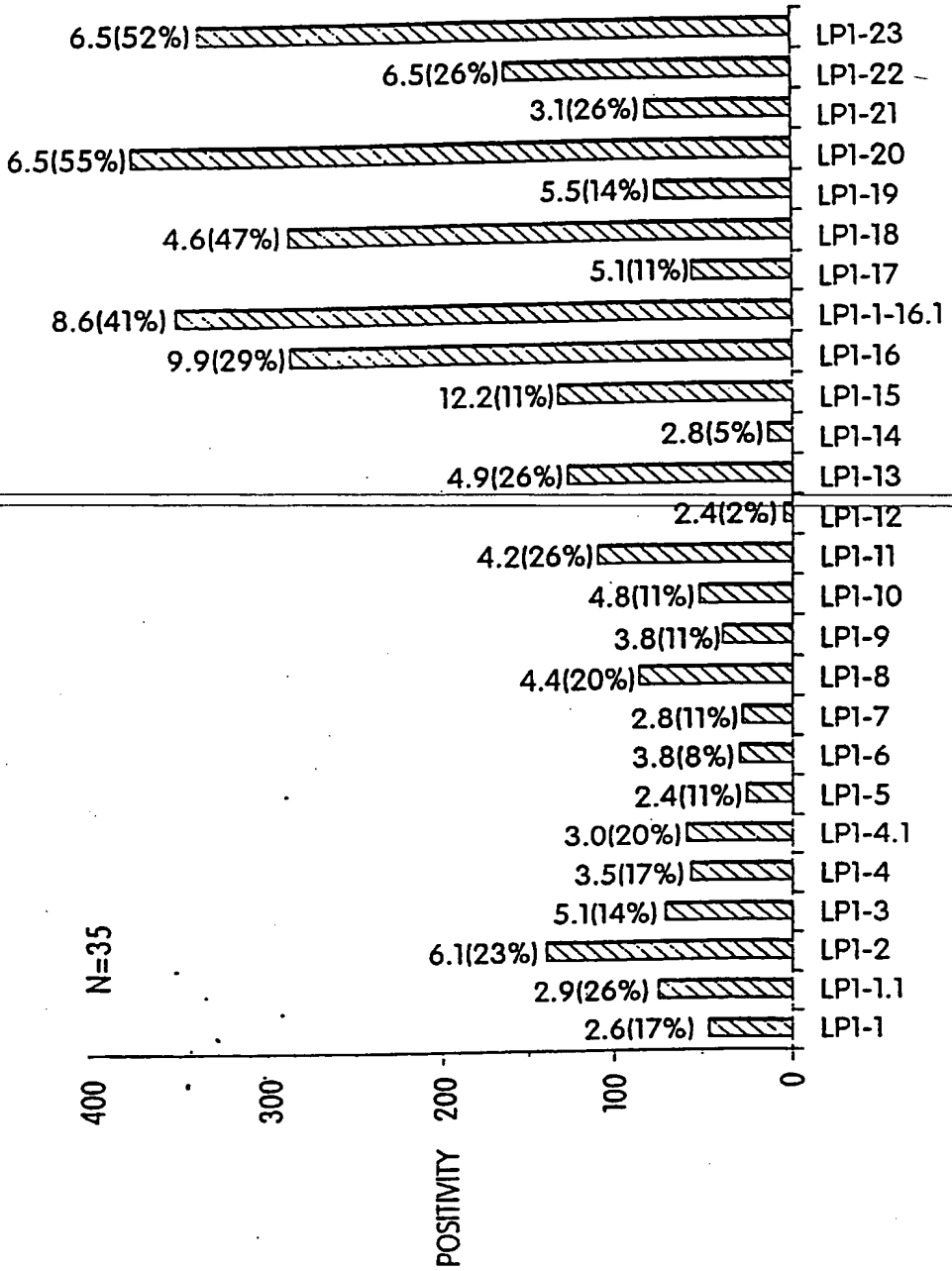

6.5(52%)
LP1-23
6.5(26%)
LP1-22
3.1(26%)
LP1-21
6.5(55%)
LP1-20
5.5(14%)
LP1-19
4.6(47%)
LP1-18
5.1(11%)
LP1-17
8.6(41%)
LP1-1-16.1
9.9(29%)
LP1-16
12.2(11%)
LP1-15
2.8(5%)
LP1-14
4.9(26%)
LP1-13
2.4(2%)
LP1-12
4.2(26%)
LP1-11
4.8(11%)
LP1-10
3.8(11%)
LP1-9
4.4(20%)
LP1-8
2.8(11%)
LP1-7
3.8(8%)
LP1-6
2.4(11%)
LP1-5
3.0(20%)
LP1-4.1
3.5(17%)
LP1-4
5.1(14%)
LP1-3
6.1(23%)
LP1-2
2.9(26%)
LP1-1.1
2.6(17%)
LP1-1
N=35
400
300
200
100
0
POSITIVITY

Fig. 3

2157596

PEPTIDE NAME	PEPTIDE SEQUENCE
LPI-16.1	VEKGSNPNYLAILVKYVDGD
LPI-16.2	DEVEKGSNPNYLAILVKYVDGD
LPI-16.3	DEAEKGSNPNYLAILVKYVDGD
LPI-16.4	KKVEKGSNPNYLAILVKK
LPI-16.5	VEKGSNPNYLAILDE
LPI-16.6	AEKGSNPNYLAILDE
LPI-16.7	DEVEKGSNPNYLAIDE
LPI-16.9	KKAEKGSNPNYLAILVKK
LPI-16.10	DEPNYLAILVKYVDE
LPI-18	GDVVAVDIKEKGKDKWIELK
LPI-18.5	GDVVAVDIKEKGKDK
LPI-18.6	VAVDIKEKGKDKWIE
LPI-18.7	AVDIKEKGKDKWIEL
LPI-18.8	DIKEKGKDKWIELK
LPI-20	ESWGAVWRIDTPDKLTGPFT
LPI-20.2	WGAVWRIDTPDKLT
LPI-20.3	GAVWRIDTPDKLTG
LPI-20.4	WRIDTPDKLTGPFT
LPI-20.5	ESWGAVWRIDTPDK
LPI-20.6	AGAVWRIDTPDKLT
LPI-23	SEVEDVIPEGWKADTSYSAK
LPI-23.1	SEVEDVIPEGWKADT
LPI-23.2	EDVIPEGWKADTSYS
LPI-23.4	IPEGWKADTSYSAK

Fig. 4

```
ATCCCGAAGGTGCCCCCGGGCCCGAACATCACGGCGACCTACGGTGACAAGTGGCTGGAC   60
 I  P  K  V  P  P  G  P  P  N  I  T  A  T  Y  G  D  K  W  L  D
 1                 10                 20

GCGAAGAGCACATGGTACGGCAAGCCGACGGGCGCCGGCCCCAAGGACAACGGCGGCGCG  120
 A  K  S  T  W  Y  G  K  P  T  G  A  G  P  K  D  N  G  G  G  A
                   30                 40

TGCGGGTACAAGGACGTGGACAAGGCGCCGTTCAACGGCATGACCGGGTGCGGCAACACC  180
 C  G  Y  K  D  V  D  K  A  P  F  N  G  M  T  G  C  G  N  T
                   50                 60

CCCATCTTCAAGGACGGGGCGCGGGTGCCGGTTCCTGCTTCGAGATCAAGTGCACGAAGCCC  240
 P  I  F  K  D  G  R  G  C  G  S  C  F  E  I  K  C  T  K  P
                   70                 80

GAGTCGTGCTCCGGCGAGGCCGTCACCGTCCACATCACCGACGACAACGAGGAGCCCATC  300
 E  S  C  S  G  E  A  V  T  V  H  I  T  D  D  N  E  E  P  I
                   90                 100

GCGCCCTACCACTTCGACCTTTCCGGCCACGCGTTCGGTTCCATGGCGAAGAGGGCGAG  360
 A  P  Y  H  F  D  L  S  G  H  A  F  G  S  M  A  K  K  G  E
                  110                 120

GAGCAGAAGCTGCGCAGCGCGGGCGAGCTGGAGCTGCAGTTTAGGCGGGTGAAGTGCAAG  420
 E  Q  K  L  R  S  A  G  E  L  E  L  Q  F  R  R  V  K  C  K
                  130                 140
```

Fig. 5

```
TACCCCGAGGGCACCAAGGTGACCTTCCACGTCGAGAAGGGTTCCAACCCCAACTACCTG  480
 Y  P  E  G  T  K  V  T  F  H  V  E  K  G  S  N  P  N  Y  L
                  150                                    160

GCGCTGCTGGTGAAGTACGTCGACGGCGACGGCGACGTGGTGGCGGTGGATATCAAGGAG  540
 A  L  L  V  K  Y  V  D  G  D  G  D  V  V  A  V  D  I  K  E
                  170                                    180

AAGGGCAAGGACAAGTGGATCGCGCTCAAGGAGTCATGGGGAGCCATCTGGAGGGTGGAC  600
 K  G  K  D  K  W  I  A  L  K  E  S  W  G  A  I  W  R  V  D
                  190                                    200

ACCCCCGACAAGCTGACGGGCCCATTCACCGTTCGCTACACCACCGAGGGAGGCACCAAG  660
 T  P  D  K  L  T  G  P  F  T  V  R  Y  T  T  E  G  G  T  K
                  210                                    220

TCCGAAGTTGAGGACGTCATCCCCGAGGGCTGGAAGGCCGACGCCAGCTACGAGTCCAAG  720
 S  E  V  E  D  V  I  P  E  G  W  K  A  D  A  S  Y  E  S  K
                  230                                    240

TGA                                                          723
```

Fig. 5 cont.

```
ATCGCGAAGGTTCCCCCCGGCCCGAACATCACGGCGACCTACGGCGACAAGTGGCTTGAC   60
 I  A  K  V  P  P  G  P  N  I  T  A  T  Y  G  D  K  W  L  D
 1           10                20

GCGAAGAGCACCTGGTACGGCAAGCCGACCGGNBCCGGTCCCAAGGACAACGGCGGCGCG  120
 A  K  S  T  W  Y  G  K  P  T  G  X  G  P  K  D  N  G  G  G  A
                30                40

TGCGGATACAAGGACGTGGACAAGCCCCCGTTCAGCGGCATGACCGGCTGCGGCAACACC  180
 C  G  Y  K  D  V  D  K  P  P  F  S  G  M  T  G  C  G  N  T
                50                60

CCCATCTTCAAGTCCGGCCGCGGCTGCGGCTCCTGCTTCGAGATCAAGTGCACCAAGCCC  240
 P  I  F  K  S  G  R  G  C  G  S  C  F  E  I  K  C  T  K  P
                70                80

GAGTCCTGCTCCGGGGAGCCCGTCCTGGTCCACATCACCGACAACGAGGAGCCCATC  300
 E  S  C  S  G  E  P  V  L  V  H  I  T  D  D  N  E  E  P  I
                90                100

GCCGCCTACCACTTCGACCTCTCCGGCAAGGCGTTCGGGGCCATGGCCAAGAAGGGTGAG  360
 A  A  Y  H  F  D  L  S  G  K  A  F  G  A  M  A  K  K  G  E
                110               120
```

Fig. 6

```
GAGCAGAAGCTGCGCAGCGCCGGCGAGCTCAAGTTCCGCCGCGGTCAAGTGCGAG  420
  E  Q  K  L  R  S  A  G  E  L  E  L  K  F  R  R  V  K  C  E
                 130                                   140

TACCCGAAGGGCACCAAGGTTACCTTCCACGTCGAGAAGGGGTCCAACCCCAACTACCTT  480
  Y  P  K  G  T  K  V  T  F  H  V  E  K  G  S  N  P  N  Y  L
                 150                                   160

GCGCTGCTGGTGAAGTACGTCGACGGCGACGGGGACGTGGTGGCGGTGGACATCAAGCAG  540
  A  L  L  V  K  Y  V  D  G  D  G  D  V  V  A  V  D  I  K  Q
                 170                                   180

AAGGGCAAGGACAAGTGGATCGAGCTCAAGGAGTCGTGGGGAGCCGTCTGGAGGATCGAC  600
  K  G  K  D  K  W  I  E  L  K  E  S  W  G  A  V  W  R  I  D
                 190                                   200

ACCCCCGACAAGCTCACCGGCCCCTTCACCGTCCGCTACACCACCGAGGGCGGCACCAAG  660
  T  P  D  K  L  T  G  P  F  T  V  R  Y  T  T  E  G  G  T  K
                 210                                   220

GCCGAAGCCGAGGACGTCATCCCCGAGGGCTGGAAGGCCGACACCGCCTACGAGGCCAAG  720
  A  E  A  E  D  V  I  P  E  G  W  K  A  D  T  A  Y  E  A  K
                 230                                   240

TGA                                                      723
```

Fig. 6 cont.

```
ATCGCGAAGGTGCCCCCGGGTCCGAACATCACGGCGACCTACGGCGACAAGTGGCTCGAC   60
 I  A  K  V  P  P  P  G  G  P  N  I  T  A  T  Y  G  D  K  W  L  D
 1           10                                      20

GCGAAGAGCACATGGTACGGCAAGCCGACGGGGGCCGGTCCCAAGGACAACGGCGGCGCT  120
 A  K  S  T  W  Y  G  K  P  T  G  A  G  P  K  D  N  G  G  G  A
                    30                                      40

TGCGGGTACAAGGACGTGGACAAGCCCCCCGTTCAGCGGCATGACCGGCTGCGGCAACACC  180
 C  G  Y  K  D  V  D  K  P  P  F  S  G  M  T  G  C  G  N  T
                 50                                      60

CCCATCTTCAAGTCCGGCCGTGGCTGCGGCTCCTGCTTTGAGATCAAGTGCACGAAGCCC  240
 P  I  F  K  S  G  R  G  C  G  S  C  F  E  I  K  C  T  K  P
                    70                                      80

GAGGCCTGCTCCGGCGAGCCCGTGGTAGTCCACATCACCGACAACGAGGAGCCCATC  300
 E  A  C  S  G  E  P  V  V  V  H  I  T  D  D  N  E  E  P  I
                 90                                     100

GCCCCCTACCACTTCGACCTCTCCGGCCACGCGTTCGGGGCGATGGCCAAGAAGGGCGAT  360
 A  P  Y  H  F  D  L  S  G  H  A  F  G  A  M  A  K  K  G  D
                   110                                     120
```

Fig. 7

```
GAGCAGAAGCTGCGCACGGCCGGCGAGCTGGAGCTCCAGTTCCGGCGCGTCAAGTGCAAG  420
  E  Q  K  L  R  R  T  A  G  G  E  L  E  L  Q  F  R  R  V  K  C  K
                                130                           140

TACCCGGAGGGGACCAAGGTGACCTTCCACGTGGAGAAGGGGTCCAACCCCAACTACCTG  480
  Y  P  E  G  T  K  V  T  F  H  V  E  K  G  S  N  P  N  Y  L
                   150                           160

GCGCTGCTTGTGAAGTACGTTAACGGCGACGGAGACGTGGTGGCGGTGGACATCAAGGAG  540
  A  L  L  V  K  Y  V  N  G  D  G  D  V  V  A  V  D  I  K  E
                   170                           180

AAGGGCAAGGACAAGTGGATCGAGCTCAAGGAGTCGTGGGGAGCCATCTGGAGGATCGAC  600
  K  G  K  D  K  W  I  E  L  K  E  S  W  G  A  I  W  R  I  D
                   190                           200

ACTCCCGACAAGCTCACGGGCCCCTTCACCGTCCGCTACACCACCGAGGGCGGCACCAAG  660
  T  P  D  K  L  T  G  P  F  T  V  R  Y  T  T  E  G  G  T  K
                   210                           220

ACCGAAGCCGAGGACGTCATCCCTGAGGGCTGGAAGGCCGACACCAGCTACGAGTCCAAG  720
  T  E  A  E  D  V  I  P  E  G  W  K  A  D  T  S  Y  E  S  K
                   230                           240

TGA                                                           723
```

Fig. 7 cont.

2157596

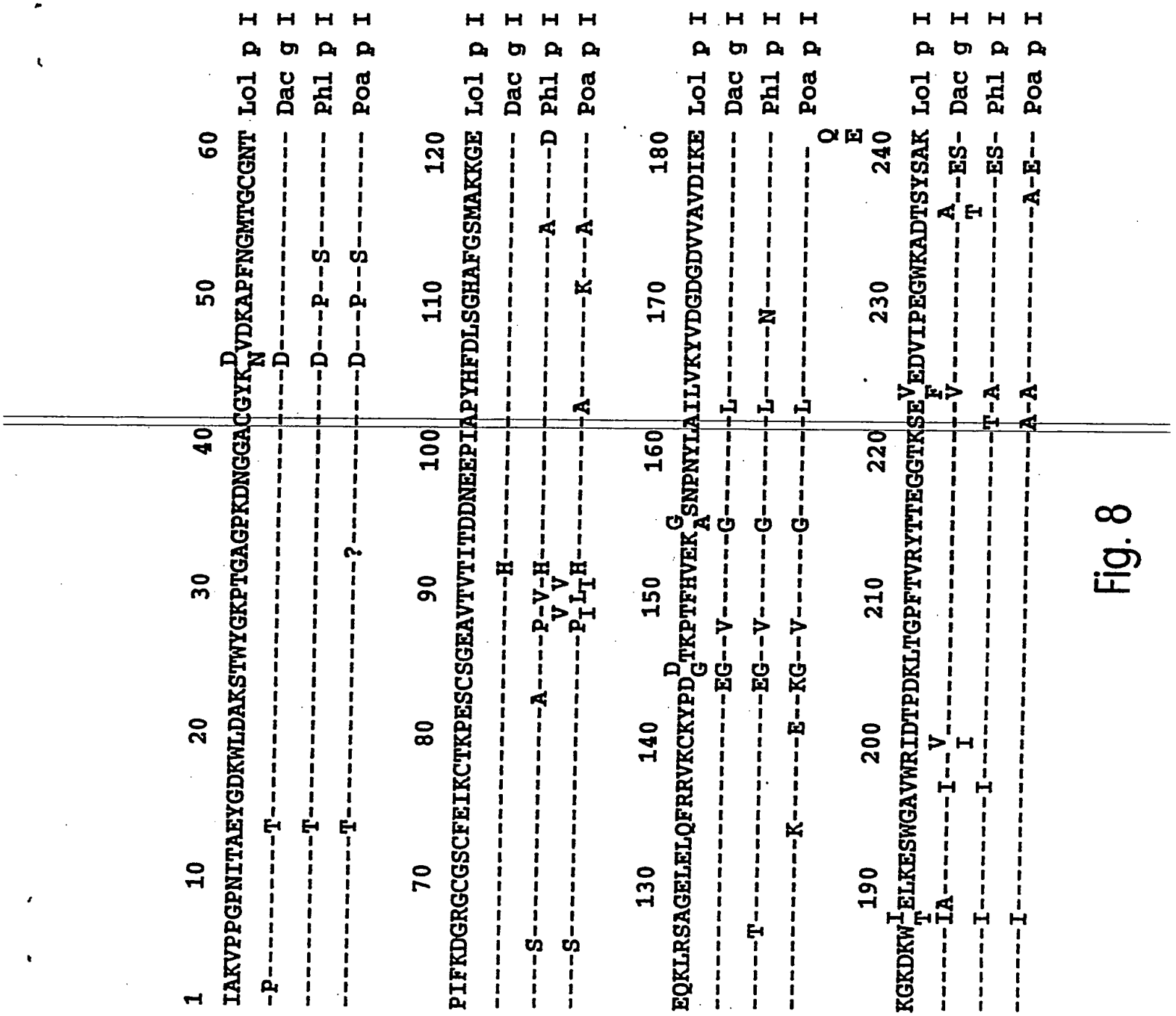
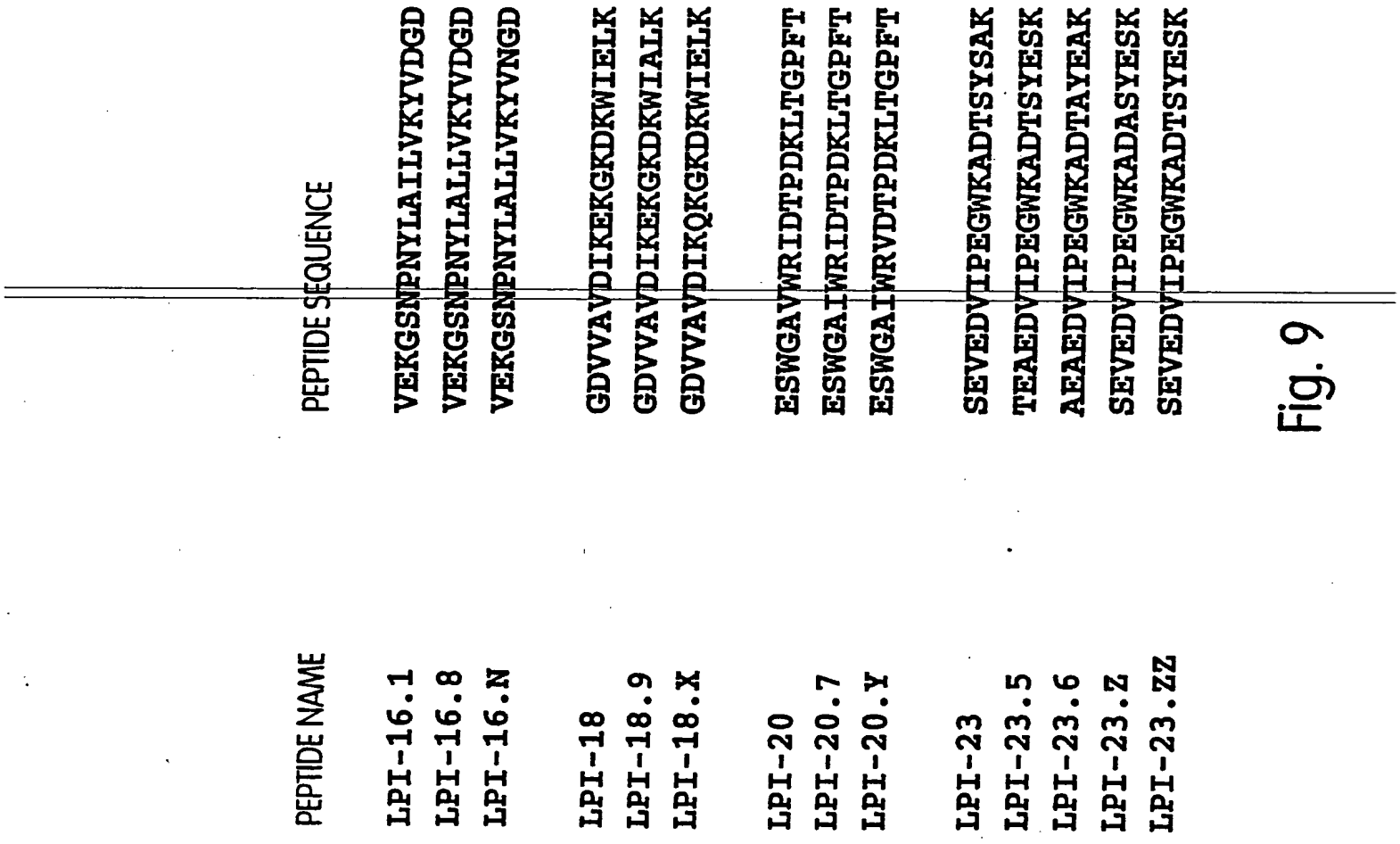


Fig. 8



PEPTIDE NAME	PEPTIDE SEQUENCE
LPI-16.1	VEKGSNPNYLAILVKYVDGD
LPI-16.8	VEKGSNPNYLALLVKYVDGD
LPI-16.N	VEKGSNPNYLALLVKYVNGD
LPI-18	GDVVAVDIKEKGKDKWIELK
LPI-18.9	GDVVAVDIKEKGKDKWIALK
LPI-18.X	GDVVAVDIKQKGKDKWIELK
LPI-20	ESWGAVWRIDTPDKLTGPFT
LPI-20.7	ESWGAIWRIDTPDKLTGPFT
LPI-20.Y	ESWGAIWRVDTPDKLTGPFT
LPI-23	SEVEDVIPEGWKADTSYSAK
LPI-23.5	TEAEDVIPEGWKADTSYESK
LPI-23.6	AEAEDVIPEGWKADTAYEAK
LPI-23.Z	SEVEDVIPEGWKADASYESK
LPI-23.ZZ	SEVEDVIPEGWKADTSYESK

Fig. 9

```
                                                                  387
GAG GCT GTC ACC GTC ACA ATC ACC GAC GAC AAC GAG GAG CCC ATC GCA
Glu Ala Val Thr Val Thr Ile Thr Asp Asp Asn Glu Glu Pro Ile Ala
     90                              95                 100

                                                                  435
CCC TAC CAT TTC GAC CTC TCG GGC CAC GCG TTC GGG TCC ATG GCG AAG
Pro Tyr His Phe Asp Leu Ser Gly His Ala Phe Gly Ser Met Ala Lys
    105                             110                 115

                                                                  483
AAG GGC GAG GAG CAG AAG CTC CGC AGC GCC GGC GAG CTG GAG CTC CAG
Lys Gly Glu Glu Gln Lys Leu Arg Ser Ala Gly Glu Leu Glu Leu Gln
    120                             125                 130

                                                                  531
TTC AGG CGG GTC AAG TGC AAG TAC CCG GAC GGC ACC AAG CCG ACA TTC
Phe Arg Arg Val Lys Cys Lys Tyr Pro Asp Gly Thr Lys Pro Thr Phe
    135                             140                 145

                                                                  579
CAC GTC GAG AAG GCT TCC AAC CCC AAC TAC CTC GCT ATT CTG GTG AAG
His Val Glu Lys Ala Ser Asn Pro Asn Tyr Leu Ala Ile Leu Val Lys
150                             155                 160                 165

                                                                  627
TAC GTC GAC GGC GAC GGT GAC GTG GTG GCG GTG GAC ATC AAG GAG AAG
Tyr Val Asp Gly Asp Gly Asp Val Val Ala Val Asp Ile Lys Glu Lys
                170                 175                 180
```

Fig. 1 cont.

2157596

```
                                                                    675
                                                GGA GCA GTC TGG
                                                Gly Ala Val Trp
                                                    195
                                                                    723
                                                ACC GTC CGC TAC
                                                Thr Val Arg Tyr
                                                    210
                                                                    771
                                                GTC ATC CCT GAG
                                                Val Ile Pro Glu
GGC AAG GAT AAG TGG ATC GAG CTC AAG GAG TCG TGG
Gly Lys Asp Lys Trp Ile Glu Leu Lys Glu Ser Trp
        185             190
AGG ATC GAC ACC CCC GAT AAG CTG ACG GGC CCA TTC
Arg Ile Asp Thr Pro Asp Lys Leu Thr Gly Pro Phe
    200             205
ACC ACC GAG GGC GGC ACC AAA TCC GAA GTC GAG GAT
Thr Thr Glu Gly Gly Thr Lys Ser Glu Val Glu Asp
215             220             225
GGC TGG AAG GCC GAC ACC TCC TAC TCG GCC AAG TGAGCA                   810
Gly Trp Lys Ala Asp Thr Ser Tyr Ser Ala Lys
230             235             240
```

Fig. 1 cont.

2157596

PEPTIDE NAME	PEPTIDE SEQUENCE
LPI-1	IAKVPPGPNITAEYGDKWLD
LPI-1.1	IAKVXPGXNITAEYGDKWLD
LPI-2	TAEYGDKWLDAKSTWYGKPT
LPI-3	AKSTWYGKPTGAGPKDNGGA
LPI-4	GAGPKDNGGACGYKNVDKAP
LPI-4.1	GAGPKDNGGACGYKDVDKAP
LPI-5	CGYKDVDKAPFNGMTGCGNT
LPI-6	FNGMTGCGNTPIFKDGRGCG
LPI-7	PIFKDGRGCGSCFEIKCTKP
LPI-8	SCFEIKCTKPESCSGEAVTV
LPI-9	ESCSGEAVTVTITDDNEEPI
LPI-10	TITDDNEEPIAPYHFDLSGH
LPI-11	APYHFDLSGHAFGSMADDGE
LPI-11.1	APYHFDLSGHAFGSMAKKGE
LPI-12	AFGSMADDGEEQKLRSAGEL
LPI-12.1	AFGSMAKKGEEQKLRSAGEL
LPI-13	EQKLRSAGELELQFRRVKCK
LPI-14	ELQFRRVKCKYPDDTKPTFH
LPI-15	YPDDTKPTFHVEKASNPNYL
LPI-16	VEKASNPNYLAILVKYVDGD
LPI-16.1	VEKGSNPNYLAILVKYVDGD
LPI-17	AILVKYVDGDGDVVAVDIKE
LPI-18	GDVVAVDIKEKGKDKWIELK
LPI-19	KGKDKWIELKESWGAVWRID
LPI-20	ESWGAVWRIDTPDKLTGPFT
LPI-21	TPDKLTGPFTVRYTTEGGTK
LPI-22	VRYTTEGGTKSEVEDVIPEG
LPI-23	SEVEDVIPEGWKADTSYSAK

Fig. 2

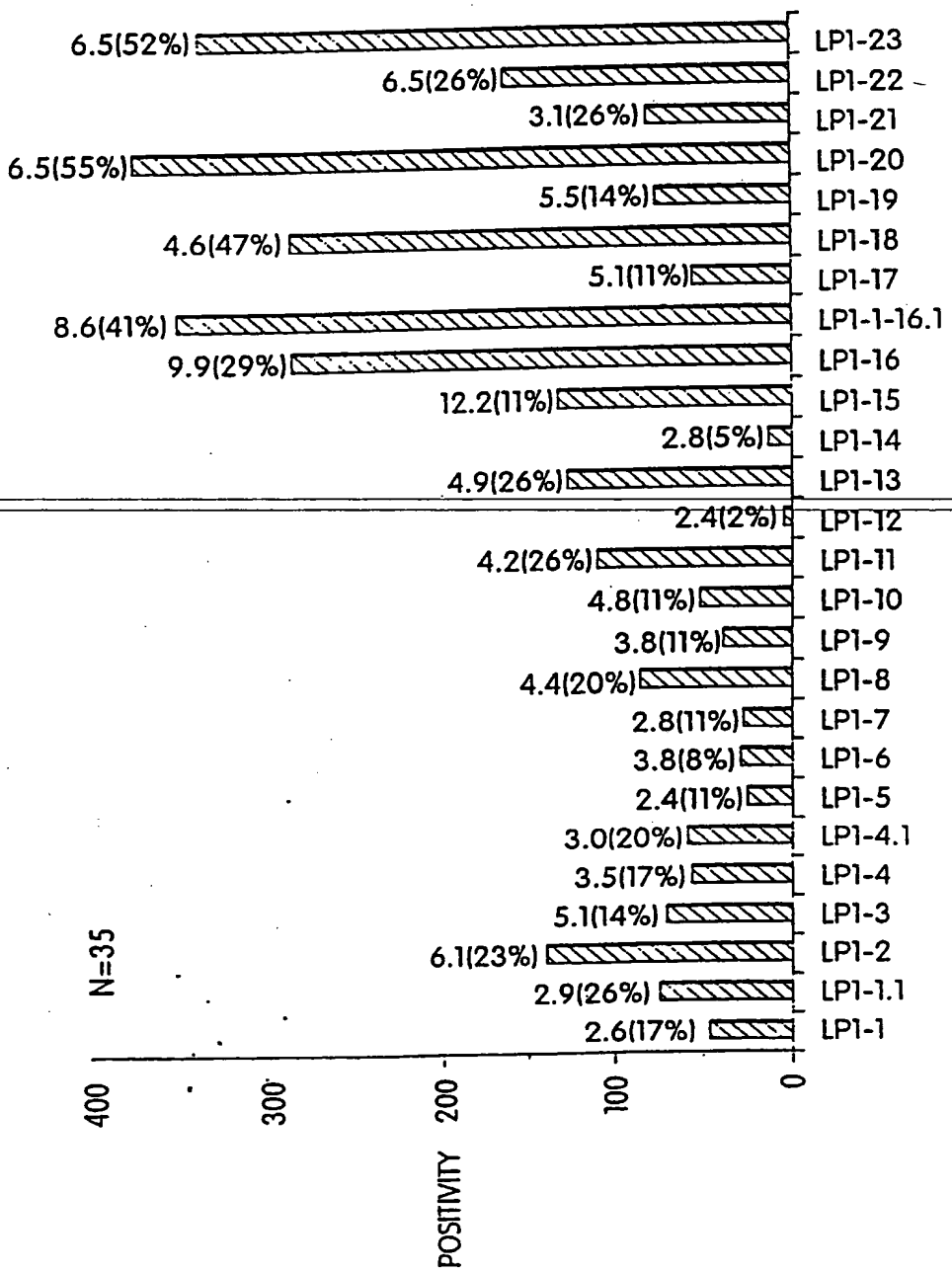

6.5(52%)
LP1-23
6.5(26%)
LP1-22
3.1(26%)
LP1-21
6.5(55%)
LP1-20
5.5(14%)
LP1-19
4.6(47%)
LP1-18
5.1(11%)
LP1-17
8.6(41%)
LP1-1-16.1
9.9(29%)
LP1-16
12.2(11%)
LP1-15
2.8(5%)
LP1-14
4.9(26%)
LP1-13
2.4(2%)
LP1-12
4.2(26%)
LP1-11
4.8(11%)
LP1-10
3.8(11%)
LP1-9
4.4(20%)
LP1-8
2.8(11%)
LP1-7
3.8(8%)
LP1-6
2.4(11%)
LP1-5
3.0(20%)
LP1-4.1
3.5(17%)
LP1-4
5.1(14%)
LP1-3
6.1(23%)
LP1-2
2.9(26%)
LP1-1.1
2.6(17%)
LP1-1
N=35
400
300
200
100
0
POSITIVITY

Fig. 3

2157596

PEPTIDE NAME	PEPTIDE SEQUENCE
LPI-16.1	VEKGSNPNYLAILVKYVDGD
LPI-16.2	DEVEKGSNPNYLAILVKYVDGD
LPI-16.3	DEAEKGSNPNYLAILVKYVDGD
LPI-16.4	KKVEKGSNPNYLAILVKK
LPI-16.5	VEKGSNPNYLAIIDE
LPI-16.6	AEKGSNPNYLAIIDE
LPI-16.7	DEVEKGSNPNYLAIDE
LPI-16.9	KKAEKGSNPNYLAILVKK
LPI-16.10	DEPNYLAILVKYVDE
LPI-18	GDVVAVDIKEKGKDKWIELK
LPI-18.5	GDVVAVDIKEKGKDK
LPI-18.6	VAVDIKEKGKDKWIE
LPI-18.7	AVDIKEKGKDKWIEL
LPI-18.8	DIKEKGKDKWIELK
LPI-20	ESWGAVWRIDTPDKLTGPFT
LPI-20.2	WGAVWRIDTPDKLT
LPI-20.3	GAVWRIDTPDKLTG
LPI-20.4	WRIDTPDKLTGPFT
LPI-20.5	ESWGAVWRIDTPDK
LPI-20.6	AGAVWRIDTPDKLT
LPI-23	SEVEDVIPEGWKADTSYSAK
LPI-23.1	SEVEDVIPEGWKADT
LPI-23.2	EDVIPEGWKADTSYS
LPI-23.4	IPEGWKADTSYSAK

Fig. 4

```
ATCCCGAAGGTGCCCCCGGGCCCGAACATCACGGCGACCTACGGTGACAAGTGGCTGGAC   60
 I  P  K  V  P  P  G  P  P  N  I  T  A  T  Y  G  D  K  W  L  D
 1                 10                 20

GCGAAGAGCACATGGTACGGCAAGCCGACGGGCGCCGGCCCCAAGGACAACGGCGGCGCG  120
 A  K  S  T  W  Y  G  K  P  T  G  A  G  P  K  D  N  G  G  G  A
                   30                 40

TGCGGGTACAAGGACGTGGACAAGGCGCCGTTCAACGGCATGACCGGGTGCGGCAACACC  180
 C  G  Y  K  D  V  D  K  A  P  F  N  G  M  T  G  C  G  N  T
                   50                 60

CCCATCTTCAAGGACGGGGCGCGGGTGCCGGTTCCTGCTTCGAGATCAAGTGCACGAAGCCC  240
 P  I  F  K  D  G  R  G  C  G  S  C  F  E  I  K  C  T  K  P
                   70                 80

GAGTCGTGCTCCGGCGAGGCCGTCACCGTCCACATCACCGACGACAACGAGGAGCCCATC  300
 E  S  C  S  G  E  A  V  T  V  H  I  T  D  D  N  E  E  P  I
                   90                100

GCGCCCTACCACTTCGACCTTTCCGGCCACGCGTTCGGTTCCATGGCGAAGAAGGGCGAG  360
 A  P  Y  H  F  D  L  S  G  H  A  F  G  S  M  A  K  K  G  E
                  110                120

GAGCAGAAGCTGCGCAGCGCGGGCGAGCTGGAGCTGCAGTTTAGGCGGGTGAAGTGCAAG  420
 E  Q  K  L  R  S  A  G  E  L  E  L  Q  F  R  R  V  K  C  K
                  130                140
```

Fig. 5

```
TACCCCGAGGGCACCAAGGTGACCTTCCACGTCGAGAAGGGTTCCAACCCCAACTACCTG  480
 Y  P  E  G  T  K  V  T  F  H  V  E  K  G  S  N  P  N  Y  L
                  150                                    160

GCGCTGCTGGTGAAGTACGTCGACGGCGACGGCGACGTGGTGGCGGTGGATATCAAGGAG  540
 A  L  L  V  K  Y  V  D  G  D  G  D  V  V  A  V  D  I  K  E
                  170                                    180

AAGGGCAAGGACAAGTGGATCGCGCTCAAGGAGTCATGGGGAGGCCATCTGGAGGGTGGAC  600
 K  G  K  D  K  W  I  A  L  K  E  S  W  G  A  I  W  R  V  D
                  190                                    200

ACCCCCGACAAGCTGACGGGCCCATTCACCGTTCGCTACACCACCGAGGGAGGCACCAAG  660
 T  P  D  K  L  T  G  P  F  T  V  R  Y  T  T  E  G  G  T  K
                  210                                    220

TCCGAAGTTGAGGACGTCATCCCCGAGGGCTGGAAGGCCGACGCCAGCTACGAGTCCAAG  720
 S  E  V  E  D  V  I  P  E  G  W  K  A  D  A  S  Y  E  S  K
                  230                                    240

TGA                                                        723
```

Fig. 5 cont.

```
ATCGCGAAGGTTCCCCCCGGCCCGAACATCACGGCGACCTACGGCGACAAGTGGCTTGAC   60
 I  A  K  K  V  P  P  G  P  P  N  I  T  A  T  Y  G  G  D  K  W  L  D
 1              10                 20

GCGAAGAGCACCTGGTACGGCAAGCCGACCGGNBCCGGTCCCAAGGACAACGGCGGCGCG  120
 A  K  S  T  W  Y  G  K  P  T  G  X  G  P  K  D  N  G  G  A
                30                 40

TGCGGATACAAGGACGTGGACAAGCCCCCCGTTCAGCGGCATGACCGGCTGCGGCAACACC  180
 C  G  Y  K  D  V  D  K  P  P  F  S  G  M  T  G  C  G  N  T
                50                 60

CCCATCTTCAAGTCCGGCCGCGGCTGCGGCTCCTGCTTCGAGATCAAGTGCACCAAGCCC  240
 P  I  F  K  S  G  R  G  C  G  S  C  F  E  I  K  C  T  K  P
                70                 80

GAGTCCTGCTCCGGGGAGCCCGTCCTGGTCCACATCACCGACGACAACGAGGAGCCCATC  300
 E  S  C  S  G  E  P  V  L  V  H  I  T  D  D  N  E  E  P  I
                90                 100

GCCGCCTACCACTTCGACCTCTCCGGCAAGGCGTTCGGGGCCATGGCCAAGAAGGGTGAG  360
 A  A  Y  H  F  D  L  S  G  K  A  F  G  A  M  A  K  K  G  E
               110                 120
```

Fig. 6

```
GAGCAGAAGCTGCGCAGCGCCGGCGAGCTGGAGCTCAAGTTCCGCCGCGTCAAGTGCGAG  420
  E  Q  K  L  R  S  A  G  E  L  E  L  K  F  R  R  V  K  C  E
                  130                                   140

TACCCGAAGGGCACCAAGGTTACCTTCCACGTCGAGAAGGGGTCCAACCCCAACTACCTT  480
  Y  P  K  G  T  K  V  T  F  H  V  E  K  G  S  N  P  N  Y  L
                  150                                   160

GCGCTGCTGGTGAAGTACGTCGACGGCGACGGGACGTGGTGCGGTGGACATCAAGCAG  540
  A  L  L  V  K  Y  V  D  G  D  G  D  V  V  A  V  D  I  K  Q
                  170                                   180

AAGGGCAAGGACAAGTGGATCGAGCTCAAGGAGTCGTGGGGAGCCGTCTGGAGGATCGAC  600
  K  G  K  D  K  W  I  E  L  K  E  S  W  G  A  V  W  R  I  D
                  190                                   200

ACCCCCGACAAGCTCACCGGCCCCTTCACCGTCCGCTACACCACCGAGGGCGGCACCAAG  660
  T  P  D  K  L  T  G  P  F  T  V  R  Y  T  T  E  G  G  T  K
                  210                                   220

GCCGAAGCCGAGGACGTCATCCCCGAGGGCTGGAAGGCCGACACCGCCTACGAGGCCAAG  720
  A  E  A  E  D  V  I  P  E  G  W  K  A  D  T  A  Y  E  A  K
                  230                                   240

TGA                                                      723
```

Fig. 6 cont.

```
ATCGCGAAGGTGCCCCCGGGTCCGAACATCACGGCGACCTACGGCGACAAGTGGCTCGAC   60
 I  A  K  V  P  P  G  G  P  N  I  T  A  T  Y  G  D  K  W  L  D
 1           10                                      20

GCGAAGAGCACATGGTACGGCAAGCCGACGGGGGCCGGTCCCAAGGACAACGGCGGCGCT  120
 A  K  S  T  W  Y  G  K  P  T  G  A  G  P  K  D  N  G  G  G  A
                    30                                      40

TGCGGGTACAAGGACGTGGACAAGCCCCCCGTTCAGCGGCATGACCGGCTGCGGCAACACC  180
 C  G  Y  K  D  V  D  K  P  P  F  S  G  M  T  G  C  G  N  T
                 50                                      60

CCCATCTTCAAGTCCGGCCGTGGCTGCGGCTCCTGCTTTGAGATCAAGTGCACGAAGCCC  240
 P  I  F  K  S  G  R  G  C  G  S  C  F  E  I  K  C  T  K  P
                    70                                      80

GAGGCCTGCTCCGGCGAGCCCGTGGTAGTCCACATCACCGACAACGAGGAGCCCATC  300
 E  A  C  S  G  E  P  V  V  V  H  I  T  D  D  N  E  E  P  I
                 90                                     100

GCCCCCTACCACTTCGACCTCTCCGGCCACGCGTTCGGGGCGATGGCCAAGAAGGGCGAT  360
 A  P  Y  H  F  D  L  S  G  H  A  F  G  A  M  A  K  K  G  D
                   110                                     120
```

Fig. 7

```
GAGCAGAAGCTGCGCACGGCCGGCGAGCTGGAGCTCCAGTTCCGGCGCGTCAAGTGCAAG  420
  E  Q  K  L  R  R  T  A  G  G  E  L  E  L  Q  F  R  R  V  K  C  K
                       130                                   140

TACCCGGAGGGGACCAAGGTGACCTTCCACGTGGAGAAGGGGTCCAACCCCAACTACCTG  480
  Y  P  E  G  T  K  V  T  F  H  V  E  K  G  S  N  P  N  Y  L
                       150                                   160

GCGCTGCTTGTGAAGTACGTTAACGGCGACGGAGACGTGGTGGCGGTGGACATCAAGGAG  540
  A  L  L  V  K  Y  V  N  G  D  G  D  V  V  A  V  D  I  K  E
                       170                                   180

AAGGGCAAGGACAAGTGGATCGAGCTCAAGGAGTCGTGGGGAGCCATCTGGAGGATCGAC  600
  K  G  K  D  K  W  I  E  L  K  E  S  W  G  A  I  W  R  I  D
                       190                                   200

ACTCCCGACAAGCTCACGGGCCCCTTCACCGTCCGCTACACCACCGAGGGCGGCACCAAG  660
  T  P  D  K  L  T  G  P  F  T  V  R  Y  T  T  E  G  G  T  K
                       210                                   220

ACCGAAGCCGAGGACGTCATCCCTGAGGGCTGGAAGGCCGACACCAGCTACGAGTCCAAG  720
  T  E  A  E  D  V  I  P  E  G  W  K  A  D  T  S  Y  E  S  K
                       230                                   240

TGA                                                           723
```

Fig. 7 cont.

```
1       10        20        30        40        50        60
                                            D
IAKVPPGPNITAEYGDKWLDAKSTWYGKPTGAGPKDNGGACGYK VDKAPFNGMTGCGNT Lol p I
                                            N
-P---------T--------------------------------D--------------- Dac g I
-----------T--------------------------------D---P--S-------- Phl p I
-----------T--------------------?-----------D---P--S-------- Poa p I

         70        80        90       100       110       120
PIFKDGRGCGSCFEIKCTKPESCSGEAVTVTITDDNEEPIAPYHFDLSGHAFGSMAKKGE Lol p I
------------------------------H----------------------------- Dac g I
----S----------------A----P-V-H---------------------A------D Phl p I
                           V V
----S----------------------P L H---------A-------K---A------- Poa p I
                           I I

         130       140       150       160       170       180
                       D         G
EQKLRSAGELELQFRRVKCKYPD TKPTFHVEK SNPNYLAILVKYVDGDGDVVAVDIKE Lol p I
                       G         A
----------------------EG--V------G-------L------------------ Dac g I
-----T----------------EG--V------G-------L-----N------------ Phl p I
------------K------E--KG--V------G-------L------------------ Poa p I
                                                           Q
                                                           E

         190       200       210       220       230       240
      I                                   V
KGKDKW ELKESWGAVWRIDTPDKLTGPFTVRYTTEGGTKSE EDVIPEGWKADTSYSAK Lol p I
      T                                   F
                  V                                    A
------IA-------I-- -----------------------V----------- --ES- Dac g I
                  I                                    T
------I--------I-------------------------T-A-------------ES- Phl p I
------I----------------------------------A-A----------A-E--- Poa p I
```

Fig. 8

PEPTIDE NAME	PEPTIDE SEQUENCE
LPI-16.1	VEKGSNPNYLALLVKYVDGD
LPI-16.8	VEKGSNPNYLALLVKYVDGD
LPI-16.N	VEKGSNPNYLALLVKYVNGD
LPI-18	GDVVAVDIKEKGKDKWIELK
LPI-18.9	GDVVAVDIKEKGKDKWIALK
LPI-18.X	GDVVAVDIKQKGKDKWIELK
LPI-20	ESWGAVWRIDTPDKLTGPFT
LPI-20.7	ESWGAIWRIDTPDKLTGPFT
LPI-20.Y	ESWGAIWRVDTPDKLTGPFT
LPI-23	SEVEDVIPEGWKADTSYSAK
LPI-23.5	TEAEDVIPEGWKADTSYESK
LPI-23.6	AEAEDVIPEGWKADTAYEAK
LPI-23.Z	SEVEDVIPEGWKADASYESK
LPI-23.ZZ	SEVEDVIPEGWKADTSYESK

Fig. 9